БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

**Выпускная работа по**

**«Основам информационных технологий»**

Магистрант

кафедры ботаники

Пинчук Вероника Владимировна

Руководители:

кандидат биологических наук, доцент Тихомиров Валерий Николаевич

старший преподаватель

Шешко Сергей Михайлович

**МИНСК - 2010**

# Оглавление:

[Оглавление: - 2 -](#_Toc250675943)

[Реферат (статья) по ИТ в предметной области – «Применение статистического пакета SPSS в ботанике» - 3 -](#_Toc250675944)

[Введение - 4 -](#_Toc250675945)

[Глава 1. Появление SPSS - революция в статистическом анализе. История развития. Описание программы. - 6 -](#_Toc250675946)

[*История появления* - 6 -](#_Toc250675947)

[*Описание программы* - 10 -](#_Toc250675948)

[Глава 2. Наиболее часто используемые модули – это должен знать каждый. Дополнительные модули. - 13 -](#_Toc250675949)

[Глава 3. Методы, применяемые для исследования. - 16 -](#_Toc250675950)

[Заключение - 23 -](#_Toc250675951)

[Список литературы к реферату - 24 -](#_Toc250675952)

[Предметный указатель к реферату - 24 -](#_Toc250675953)

[Интернет ресурсы в предметной области исследования - 25 -](#_Toc250675954)

[Действующий личный сайт в WWW - 28 -](#_Toc250675955)

[Презентация магистерской диссертации - 28 -](#_Toc250675956)

[Тестовые вопросы по Основам информационных технологий - 28 -](#_Toc250675957)

[Граф научных интересов - 29 -](#_Toc250675958)

[Приложение - 30 -](#_Toc250675959)

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

# Реферат (статья) по ИТ в предметной области – «Применение статистического пакета SPSS в ботанике»

Магистрант

кафедры ботаники

Пинчук Вероника Владимировна

Руководители:

Кандидат биологических наук, доцент Тихомиров Валерий Николаевич

старший преподаватель

Шешко Сергей Михайлович

**МИНСК - 2010**

# Введение

Внедрение компьютеров буквально во все сферы человеческой деятельности является на сегодняшний день, наверное, самым наглядным результатом научного прогресса. И, как и следовало ожидать, в существенной степени компьютеризация изменила исследовательский процесс. Компьютер сейчас используется для выполнения такой работы, которая раньше считалась не только самой скучной и утомительной, но и рутинной – самой объемной во всех статистических исследованиях: учет и организация исходных данных, вычисление различных показателей и многие другие статистически-математические операции, сложные для понимания и вычисления. С появлением многочисленных статистических и вычислительных программ исследователи получили долгожданную возможность быстро и легко справляться с различной вычислительной работай и больше времени уделять анализу и интерпретации полученных данных.

Для статистического анализа необходима обработка очень большого объема данных, и часто сводные таблицы данных могут достигать практически необозримых размеров [1].

Многофункциональные пакеты для обработки и анализа данных, такие как Excel, Access, STADIA, STATA, JMR (StatView), SAS (компания SAS Institute), STATGRAPICS, STATISTICA, SPSS и бесчисленное множество других нашли самое широкое распространение в самых разных областях исследований. Одним из самых незаменимых помошников для исследователя, работающего в области статистического анализа данных, является программа SPSS - очень мощный и широко распространенный пакет профессионального компьютерного анализа данных [8].

SPSS - Statical Packege for the Sosial Sciance - статистический пакет для социальных наук – выполнен очень продуманно и на таком высоком уровне, что сразу получил не только широкое распространение, но и признание со стороны статистов естественных наук. Включая огромное количество различных функций и возможностей – его ресурсы практически неисчерпаемы.

За очень короткое время SPSS позволяет сделать разносторонний анализ данных [1, 2, 3].

# Глава 1. Появление SPSS - революция в статистическом анализе. История развития. Описание программы.

История появления

Два студента Норман Най (Norman Nie) и Дейл Вент (Dale Bent), специализировавшиеся в области политологии в 1965 году пытались отыскать в Стенфордском университете Сан-Франциско компьютерную программу, подходящую для анализа статистической информации. Вскоре они разочаровались в своих попытках, так как имеющиеся программы оказывались более или менее непригодными, неудачно построенными или не обеспечивали наглядность представления обработанной информации. К тому же принципы пользования менялись от программы к программе.

Так, не долго думая, они решили разработать собственную программу, со своей концепцией и единым синтаксисом. В их распоряжении тогда был язык программирования FORTRAN и вычислительная машина типа IBM 7090. Уже через год была разработана первая версия программы, которая, еще через год, в 1967, могла работать на IBM 360. К этому времени к группе разработчиков присоединился Хэдлай Халл (Hadlai Hull).

Как известно из истории развития информатики, программы тогда представляли собой пакеты перфокарт. Как раз на это указывает и исходное название программы, которое авторы дали своему продукту: SPSS - это аббревиатура от Statistical Package for the Social Science.

В 1970 году работа над программой была продолжена в Чикагском университете, а Норман Най основал соответствующую фирму - к тому моменту уже было произведено шестьдесят инсталляций. Первое руководство для пользователей (издательство McGraw-Hill) описывало одиннадцать различных процедур.

Спустя пять лет проект выделился в отдельную компанию SPSS Inc. Программа SPSS была уже инсталлирована шестьсот раз, причём под разными операционными системами. С самого начала версиям программы присваивали соответствующие порядковые номера. В 1975 была разработана уже шестая версия (SPSS6). До 1981 последовали версии 7, 8 и 9.

Командный язык (синтаксис) SPSS в то время был ещё не так хорошо развит, как сейчас, и естественно ориентирован на перфокарты. Поэтому так называемые управляющие карты SPSS состояли из идентификационного поля (столбцы 1-15) и из поля параметров (столбцы 16-80).

В 1983 году командный язык SPSS был полностью переработан, синтаксис стал значительно удобней. Что бы отметить этот факт, программа была переименована в SPSSX, где буква X должна была служить как номером версии в римскими числами, так и сокращением для extended (расширенный).

Так как применение перфокарт к этому моменту уже стало историей, то программа SPSS и информация, подлежащая обработке, сохранялись в отдельных файлах на винчестерах больших ЭВМ, которые тогда использовались повсеместно. Год от года постоянно увеличивалось и количество процедур.

С появлением персональных компьютеров была разработана также и PC-версия SPSS, с 1983 года появилась PC-версия **SPSS\PC+.** **рассчитанная на MS-DOS.** Позже, с момента основания в 1984 году европейского торгового представительства в Горинхеме в Нидерландах, SPSS стал широко применяться и в Европе. В настоящее время это самое распространённое программное обеспечение для статистического анализа во всём мире.

Для того чтобы отразить возможность использования программы во всех областях, имеющих отношение к статистическому анализу, буква X вновь была удалена из названия марки, а исходной аббревиатуре присвоено новое значение: Superior Performance Software System (система программного обеспечения высшей производительности).

Если PC версия SPSS/PC+ была чуть усовершенствованной версией для больших ЭВМ, то SPSS для операционной системой Windows (**SPSS for Windows**) стала большим шагом вперёд. Во-первых, эта версия SPSS обладает всеми возможностями версии для больших ЭВМ, во-вторых, за некоторыми немногочисленными исключениями, программой можно пользоваться без особых знаний в области прикладного программирования. Вызов необходимых процедур статистического анализа происходит при помощи стандартной техники, применяемой в Windows, то есть с помощью мыши и соответствующих диалоговых окон.

Первая версия SPSS для Windows имевшая порядковый номер 5 (**SPSS 5.0 *for Windows***) вышла в 1992 году. Затем последовали версии 6.0 и 6.1 с некоторыми нововведениями в статистической и графической областях; версия 6.1 была первой статистической программой для Windows, которая использовала 32 битную архитектуру Windows 3.1. Это можно было заметить по более высокой скорости выполнения вычислений. Усовершенствования коснулись также и интерфейса пользователя. В конце концов, была выпущена версия 6.1.3, которая уже могла работать и под Windows 95 и под NT.

В начале 1996 года появилась 7-я версия SPSS, сначала как версия 7.0, а затем 7.5. Наряду с расширением возможностей в сфере статистики, разница между этими двумя версиями заключалась в том, что в версии 7.5 как меню, так и интерфейс программы были выполнены уже не только на английском, но и на других наиболее распространенных языках.

Самым весомым отличием версии 7 по отношению к предыдущим версиям, был абсолютно новый подход к выводу информации на экран. Так, во-первых, получил новые очертания так называемый Viewer (Окно просмотра), и, во-вторых, более приятный внешний вид приобрели таблицы результатов расчётов (мобильные таблицы). Появившаяся технология мобильных таблиц позволяет перестраивать полученные таблицы различным способами.

Если предшественница данной версии - версия 6.1.3 могла работать как под старой Windows 3.1 так и под новой Windows 95 (NT), то SPSS версии 7 могла работать только при наличии Windows 95 (NT).

За версией 7.5 последовала версия 8.0, прогресс которой заключался в усовершенствовании графической оболочки. Возможность составления интерактивных графиков предоставляет рад преимуществ по сравнению с традиционными графиками, которые являются стандартом для многих других пакетов.

Версия 9.0 включала в себя несколько новых статистических методов, в т.ч. многозначную логистическую регрессию, и несколько новых графических возможностей, расширяющих область интерактивных графиков.

**Версия 10.0 пакета SPSS for Windows** – вышла в сентябре 1999 года. Начиная с версии 10, предел по числу переменных был практически снят, единственный оставшийся «жесткий» предел составляет 2.15 миллиарда переменных. [5]

**Версия 12.0 пакета SPSS for Windows** - появление возможности русификации окон вывода результатов.

**Версия 12.5 пакета SPSS for Windows** – русификация основных диалоговых окон.

**Версия 13.0 пакета SPSS for Windows** более совершенные графические возможности программы.

**Версия 14.0 пакета SPSS for Windows** – русификация основных диалоговых окон.

**Версия 15.0 пакета SPSS for Windows** - распространяемая с сентября 2006 года.

**Версия 17.0 пакета SPSS for Windows** – вышла в 2008 году. Среди основных новшеств SPSS Statistics 17.0: Возможность работать под Windows Vista (а также под Windows XP, Mac OS X и Linux). Усовершенствованный Редактор синтаксиса. Поддерживает 10 языков: английский, русский, французский, немецкий, испанский, итальянский, японский, польский, корейский, китайский. Улучшен Конструктор диаграмм и существенно расширенные возможности визуализации данных. Процедура Оптимальной категоризации (в SPSS Data Preparation).Нейронные сети (в SPSS Neural Networks). Анализ RFM (в SPSS EZ RFM). Процедуры Обобщенная линейная модель и Обобщенные уравнения (в SPSS Advanced Statistics). Процедура анализа методом ближайшего соседа. Поддержка языка скриптов Python. А также другие важные новшества и усовершенствования [11; 9].

**Версия 18.0 пакета SPSS for Windows – готовиться к выходу!**

Описание программы

SPSS для Windows (аббревиатура SPSS расшифровывается как Statistical Package for Social Science – Статистический пакет для социальных наук) - это модульный, полностью интегрированный, обладающий всеми необходимыми возможностями программный комплекс, охватывающий все этапы аналитического процесса: планирование, сбор данных, доступ к данным и управление данными, анализ, создание отчетов и распространение результатов. SPSS для Windows - это лучшее программное обеспечение, позволяющее решать бизнес-проблемы и исследовательские задачи, используя статистические методы. Графический пользовательский интерфейс упрощает работу даже, несмотря на то, что SPSS для Windows включает в себя все функции управления данными, статистические процедуры и средства создания отчетов, необходимые для самого сложного анализа. Ближайшим аналогом SPSS является программа Statistika [4].

SPSS представляет собой набор различных программ обработки данных - эти программы облегчают процесс ввода данных (информации), легко позволяет гибко менять структуру данных, использовать самые современные методы обработки данных (как по отдельности, так и совместно) и получать результаты в удобной и наглядной форме.

Все эти многочисленные программы собраны в единую цельную систему, которая обеспечивает легкость общения пользователя с достаточно сложной, комплексной программой, снабженной исчерпывающей справочной поддержкой.

SPSS - компьютерная программа для статистической обработки данных, один из лидеров рынка в области коммерческих статистических продуктов, предназначенных для проведения прикладных исследований в социальных науках. По мнению некоторых авторов, SPSS "занимает ведущее положение среди программ, предназначенных для статистической обработки информации" [11].

Основным достоинством программного комплекса SPSS, как одного из самых существенных достижений в области компьютеризированного анализа данных, является самый широкий охват существующих статистических методов, который удачно сочетается с большим количеством удобных средств визуализации результатов обработки.

Основные функциональные возможности программы по обработке статистических данных и презентации результатов: кодирование статистических данных; основные определения и понятия из области статистики: шкалы статистических измерений, выборки и распределения, статистическая значимость понятие о зависимых и независимых выборках; процедура проведения частотного анализа данных; первичная обработка и подготовка данных для применения основных статистических методов, вычислительные возможности SPSS; условный отбор данных для повышения эффективности использования процедур статистической обработки; приемы поиска ошибок при вводе информации; статистические характеристики данных и основные типы распределений, которым подчиняются выборки; проведение тестов на наличие нормального распределения; построение таблиц сопряженности и расчет статистических характеристик для них; интерпретация таблиц сопряженности и ее практическое значение; методы анализа множественных вариантов ответов и условия их применения; множественные варианты ответов и частотный анализ; множественные варианты ответов и таблицы сопряженности; проведение сравнения средних значений выборок и интерпретация результатов; проведение непараметрических тестов для выборок, которые не подчиняются нормальному распределению или не относятся к порядковой шкале; виды корреляционного анализа и интерпретация результатов; простая линейная регрессия (пример из маркетинга); многомерная регрессия (пример из маркетинга); нелинейная регрессия (пример из социологии); логистическая регрессия (пример из медицины); порядковая регрессия (пример из психологии поведения); дискриминационный анализ и его практическое применение; надежность результатов дискриминационного анализа с практической точки зрения; факторный анализ и его практическое применение, методы факторизации переменных; пример применения факторизации в маркетинге, социологии и психологии; кластерный анализ, принципы его проведения и практическое применение; иерархический кластерный анализ для двух и более переменных; кластерный анализ для большого объема данных; интерпретация результатов кластерного анализа; примеры проведения кластерного анализа в маркетинге, социологии и экономической географии; анализ пригодности заданий для статистических обследований и опросов; управление выводом таблиц в SPSS; графические презентационные возможности SPSS; экспорт данных и результатов расчетов SPSS в Microsoft Office, преобразование в HTML-формат; программный синтаксис SPSS 10/11; язык сценариев SPSS 10/11; новое в версии SPSS 11; таблица соответствий пунктов меню русской и английской версии SPSS 10 [7].

# Глава 2. Наиболее часто используемые модули – это должен знать каждый. Дополнительные модули.

**SPSS Base (Основной системный модуль,** **Базовый модуль)** входит в базовую поставку. Он включает все процедуры ввода, отбора и корректировки данных, а также большинство предлагаемых в SPSS статистических методов. Наряду с простыми методиками статистического анализа, такими как частотный анализ, расчет статистических характеристик, таблиц сопряженности, корреляций, построения графиков, этот модуль включает t-тесты и большое количество других непараметрических тестов, а также усложненные методы, такие как многомерный линейный регрессионный анализ, DA, факторный анализ, кластерный анализ, дисперсионный анализ, анализ пригодности (анализ надежности) и многомерное шкалирование.

**Regression Models (Модуль регрессионных моделей)** модуль включает в себя различные методы регрессионного анализа, такие как: бинарная и мультиномиальная логистическая регрессия, нелинейная регрессия и пробит-анализ.

**Advanced Models (Модуль дополнительных моделей)** входят различные методы дисперсионного анализа (многомерный, с учетом повторных измерений), общая линейная модель, анализ выживания, включая метод Каплана-Майера и регрессию Кокса, логлинейные, а также логитлоглинсйные модели.

Модуль **Tables** служит для создания презентационных таблиц. Здесь предоставляются более широкие возможности по сравнению с упрощенными частотными таблицами и таблицами сопряженности, которые строятся в SPSS Base (базовом модуле).

Для расширения возможностей SPSS предлагаются следующие модули и программы.

Наряду с тремя упомянутыми, существует еще ряд специальных дополнительных модулей и самостоятельных программ, число которых постоянно растёт, так, что пользователям следует постоянно знакомиться с информацией о нововведениях в SPSS.**Amos (Analysis of moment structures - анализ моментных структур)** включает методы анализа с помощью линейных структурных уравнений. Целью программы является проверка сложных теоретических связей между различными признаками случайного процесса и их описание при помощи подходящих коэффициентов. Проверка проводится в форме причинного анализа и анализа траектории. При этом пользователь в графическом виде должен задать теоретическую модель, в которую вместе с данными непосредственных наблюдений могут быть включены и так называемые скрытые элементы. Программа Amos включена в состав модулей расширения SPSS, как преемник L1SREL (Linear Structural RELationships - линейные структурные взаимоотношения).

**AnswerTree (дерево решений)** включает четыре различных метода автоматизированного деления данных на отдельные группы (сегменты). Деление проводится таким образом, что частотные распределения целевой (зависимой) переменной в различных сегментах значимо различаются. Типичным примером применения данною метода является создание характерных профилей покупателей при исследовании потребительского рынка. AnswerTree является преемницей программы СНАШ (Chi squared interaction Detector - детектор взаимодействий на основе хи-квадрата).

**Categories** модуль содержит различные методы для анализа категориальных данных, а именно: анализ соответствий и три различных метода оптимального шкалирования (анализ однородности, нелинейный анализ главных компонент, нелинейный канонический корреляционный анализ).

**Clementine** - это программа для data mining (добычи знаний), в которой пользователю предлагаются многочисленные подходы к построению моделей, к примеру, нейронные сети, деревья решений, различные виды регрессионного анализа. Clementine представляет собой "верстак" аналитика, при помощи которого можно визуализировать процесс моделирования, перепроверять модели, сравнивать их между собой. Для удобства пользования программой имеется вспомогательная среда внедрения результатов.

**Conjoint (совместный анализ)** анализ применяется при исследовании рынка для изучения потребительских свойств продуктов на предмет их привлекательности. При этом опрашиваемые респонденты по своему усмотрению должны расположить предлагаемые наборы потребительских свойств продуктов в порядке предпочтения, на основании которого можно затем вывести так называемые детализированные показатели полезности отдельных категорий каждого потребительских свойства.

Программа **Data Entry** предназначена для быстрого составления вопросников, а также ввода и чистки данных. Заданные на этапе создания вопросника вопросы и категории ответов потом используются в качестве меток переменных и значений.

**Exact Tests (Точные тесты)** данный модуль служит для вычисления точного значения вероятности ошибки (величины **ρ**) в условиях ограниченности данных при проверке по критерию **χ2** (Chi-Quadrat-Test) и при непараметрических тестах. В случае необходимости для этого также может быть применён метод Монте-Карло (Monte-Carlo).

**GOLDMineR** программа содержит специальную регрессионную модель для регрессионного анализа упорядоченных зависимых и независимых переменных.

При помощи **SamplePower** может быть определён оптимальный размер выборки для большинства методов статистического анализа, реализованных в SPSS.

**Missing Value Analysis** данный модуль служит для анализа и восстановления закономерностей, которым подчиняются пропущенные значения. Он предоставляет различные варианты замены недостающих значений.

Модуль **Trends** содержит различные методы для анализа временных рядов, такие как: модели ARIMA, экспоненциальное сглаживание, сезонная декомпозиция и спектральный анализ [10].

**Complex Samples** позволяет учитывать сложные планы выборок при проведении анализа данных опросов и обследований, а также планировать сложные выборки и производить отбор. SPSS Complex Samples позволяет существенно снизить вероятность сделать ошибочные выводы для стратифицированных, кластеризованных и многоэтапных выборок.

**Neural Networks** cодержит нелинейные процедуры моделирования, позволяющие обнаруживать более сложные взаимосвязи в данных. Процедуры этого модуля являются дополнением традиционных статистических методов, содержащихся в SPSS Statistics Base и дополнительных модулях к нему.

**Forecasting** прогнозирование временных рядов: модели экспоненциального сглаживания, а также методы оценивания авторегрессионных моделей [6].

# Глава 3. Методы, применяемые для исследования.

ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ – **One-way ANOVA** – статистический анализ, устанавливающий статистическую значимость различий между средними значениями для трех или более выборок.

Процедура сравнения средних значений выборок, на основании которой можно сделать выводы о соотношении средних значений генеральных совокупностей.

КЛАСТЕРНЫЙ АНАЛИЗ – **Cluster analysis** - процедура, на основе заданного правила объединяющая объекты или переменные в группы, называемые кластерами. Используется как «разведывательный» анализ в самостоятельном использовании не представляющий статистической ценности, однако позволяющий судить и направлении исследования (сокращает время поиска и подбора методов для каждого конкретно взятого случая исследования).

ОПИСАТЕЛЬНАЯ СТАТИСТИКА – **Descriptive statistics** – это различные вычисляемые показатели, характеризующие распределение значений переменной. Условно разделены на две группы, первая – меры центральной тенденции, вокруг которых «группируются» данные (среднее значение, медиана и мода), вторая группа характеризует изменчивость значений переменной относительно среднего: стандартное отклонение и дисперсия. Диапазон изменчивости характеризуется минимумом, максимумом и размахом. Асимметрия и эксцесс представляют меру отклонения формы распределения от нормального вида. Существуют величины, отражающие погрешности некоторых статистик.

ДИСКРИМИНАНТНЫЙ АНАЛИЗ –**DA** – процедура создания формулы регрессии, на основе которой производится разбиение объектов на группы, соответствующие категориям зависимой переменной.

Дискриминантный анализ представляет собой альтернативу множественного регрессионного анализа. Для случая, когда зависимая переменная представляет собой не количественную, а номинальную переменную. Дискриминантный анализ, по сути, решает следующие задачи: предсказание значения «зависимой» переменной и определение того, какие «независимые» переменные лучше всего подходят для такого предсказания. Дискриминантный анализ основан на составлении уравнения регрессии, использующего номинальную зависимую переменную. Уравнение регрессии составляется на основе тех объектов, о которых известна групповая принадлежность, что позволяет максимально точно подобрать его коэффициенты. После того как уравнение регрессии получено, его можно использовать для группировки интересующих исследователя объектов, в целях прогнозирования их принадлежности к какому-либо классу.

Команда дискриминантный анализ непроста и требует настройки множества параметров, работа с которыми лежит в области статистики, но также приводится в руководстве пользователя SPSS.

Параметры дискриминантного анализа определяются особенностями данных и задачами исследователя ввиду того, что дискриминантный анализ - сложная статистическая операция.

Для демонстрации дискриминантного анализа приведем пример пошаговой работы с программой, у опытных исследователей этот процесс с помощью программы занимает порядка 10-15 минут, в то время как при расчете всех этик коэффициентов и параметров «вручную» может занять до нескольких недель.

Таблицы исходных данных могут занимать очень большие объемы.

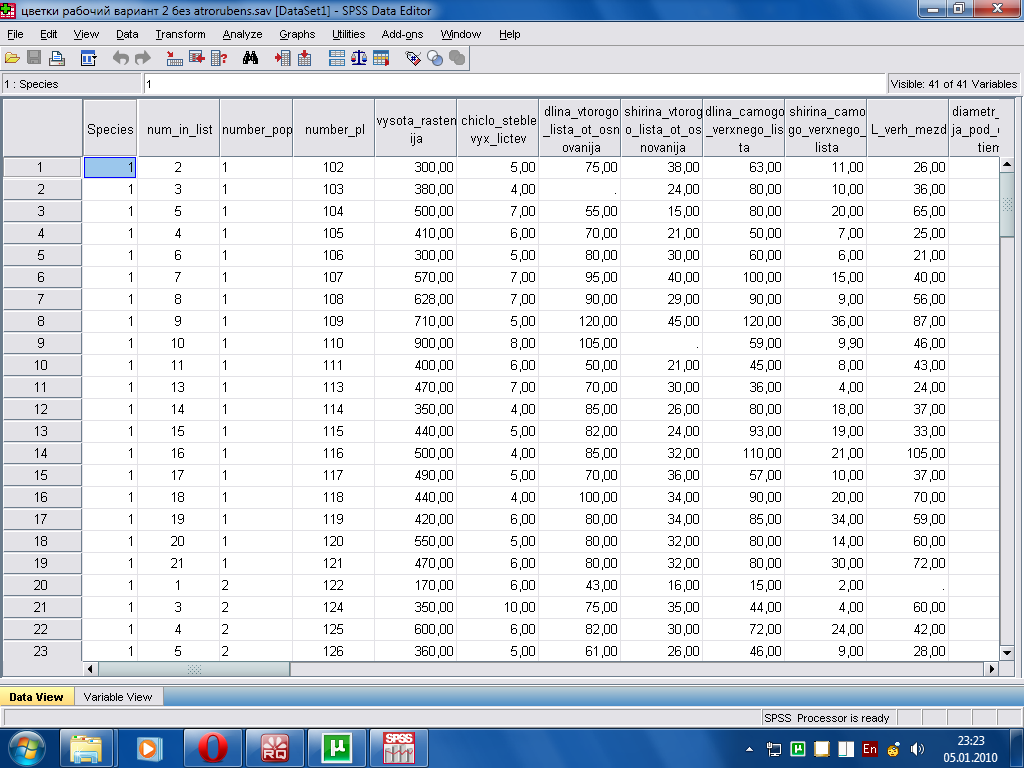


Рисунок 1. Таблица исходных данных.

Для дискриминантного анализа меню Analize (Анализ) выбираем команду Classify ► Discriminant (Классификация ► Дискриминантный анализ). На экране появляется диалоговое окно Discriminant Analisis (дискриминантный анализ), рис 2.

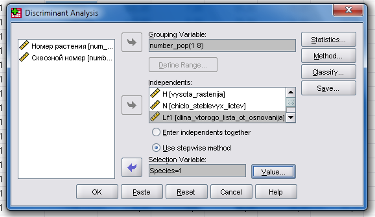


Рисунок 2. Диалоговое окно дискриминантный анализ.

Диалоговое окно содержит следующие элементы:

* Список доступных переменных (слева);
* Пять стандартных кнопок (справа) - OK, Paste, Reset, Cancel, Help;
* Четыре дополнительных элемента:

**Grouping Variable** (Группирующая переменная) для задания единственной зависимой переменной. Включает параметр Define Range (Задать диапазон) имеющий два уровня **Minumum** (Минимум)& **Maximum** (Максимум);

В список **Independents** (Независимые переменные) включается любое число переменных, участвующих в дискриминантном анализе в качестве предикторов.

Переключатели **Enter independents together** (Вводить независимые переменные вместе) и **Use stepwise method** (Использовать пошаговый метод) позволяют выбрать один из двух вариантов дискриминантного анализа.

**Selection Variable** (Переменная отбора) позволяет указать имя переменной и ее уровень. Включает параметр **Value** (Значение).

Четыре кнопки управления дополнительными параметрами: **Statistics, Method, Classify, Save.**

Кнопка **Statistics** (Статистики) открывает диалоговое окно **Discriminant Analisis: Statistics** (Дискриминантный анализ: Статистики), рис. 3. С помощью этих элементов исследователи выясняют наличие зависимостей между предикторами перед началом дискриминантного анализа. Содержит три группы: **Descriptives** (Описательные статистики): **Means** (Средние), **Univariate** **ANOVAs** (Однофакторный дисперсионный анализ), **Box’s** **M** (М Бокса); **Function Coefficient** (Коэффициенты функций): **Fisher’s** (Фишер), **Unstandardized** (Нестандартизированные коэффициенты); **Matrices** (Матрицы): **Within-group correlation** (Корреляция внутри групп), **Within-group covariance** (Ковариация внутри групп), **Separate-groups covariance** (Ковариация отдельных групп), **Total Covariance** (Полная ковариация).

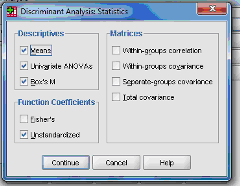


Рисунок 3. Диалоговое окно дискриминантный анализ Statistiks.

Кнопка **Method** (Метод), рис. 4, открывает диалоговое окно **Discriminant Analisis: Stepwise Method** (Дискриминантный анализ: Пошаговый метод) позволяет выбрать один из пошаговых методов составления дискриминантного уравнения (группа Method), задать критерии для включения в дискриминантное уравнение и исключение из него предикторов (группа **Criteria**), а также включить в окно вывода желаемые величины – флажки в группе **Display**.

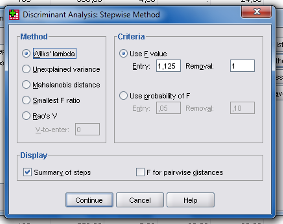


Рисунок 4. Диалоговое окно дискриминантный анализ Method.

Кнопка **Classify** (Классификация) открывает диалоговое окно **Discriminant Analisis: Classification** (Дискриминантный анализ: Классификация). Содержит четыре группы: **Prior probabilities** (Априорные вероятности); **Plots** (Диаграммы); **Use** **Covariance Matrix** (Использование ковариационной матрицы); **Display** (Отображать); **Replace Missing Values within mean** (заменять пропущенные значения средним).

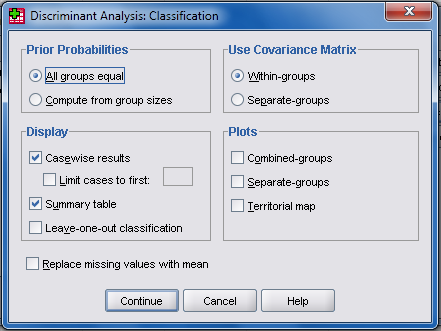


Рисунок 5. Диалоговое окно дискриминантный анализ Classification.

Кнопка **Save** (Сохранение) позволяет сохранить в качестве новых переменных следующие величины для каждого объекта:

* Прогнозируемый номер группы;
* Оценки дискриминантных функций;
* Вероятность принадлежности к каждой группе.

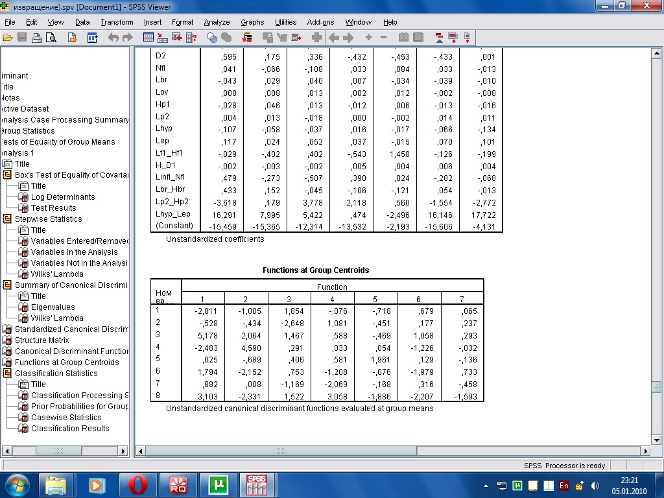


Рисунок 6. Окно вывода результатов.

После задания необходимых условий дискриминантного анализа и нажатия кнопки **OK** открывается окно вывода результатов, рис. 6. В зависимости от заданных условий исследования, программа дает ответы в таблицах, схемах и графиках. Окно вывода результатов оснащена иерархической структурой, которая делает удобной навигацию внутри окна вывода. Окно вывода результатов имеет собственную панель управления. Пользователь имеет возможность корректировать полученные результаты для удобства их восприятия, а именно**: удалять ненужные фрагменты полученных данных, добавлять свои комментарии, реорганизовывать порядок вывода, сохранять результаты в файле, менять оформление, копировать фрагменты в текстовые редакторы.** Следует обратить внимание, что окно вывода результатов практически является самостоятельной программой, что напоминает нам о том, что SPSS это многокомпонентный модуль [1, 2, 3].

# Заключение

Программа SPSS это невероятно полезный и удобный статистический пакет, который позволяет исследователю не только экономить массу времени благодаря:

1. автоматизации подсчетов,
2. комплексности программы,
3. точности математических вычислений,
4. способности дополнить недостающие данные, *но и* …
5. быстроте и удобстве получения и интерпретации результатов,
6. автоматическое построение графиков, схем, таблиц,
7. возможности создания презентаций,
8. совместимости с программами Microsoft Office.

Взамен своей полезности, удобству использования, быстроте реагирования программа требует от исследователя не только умения с ней работать, но и владения, знания и понимания статистической базы. Более высокими по сравнению с другими программами является и системные требования. Ввиду своей сложности, многокомпонентности и большого количества выполняемых математических результатов сразу программа не работает с мгновенной, молниеносной скоростью, а требует дать ей «подумать» правда, совсем немного.

За долгие годы существования программа все усложнялась и совершенствовалась, в ожидаемом в скором времени 18 выпуске пользователь получит самый совершенный статистический пакет. Помимо пополнения арсенала статистических возможностей, выйдет долгожданная полная русская версия, что значительно облегчит диалог пользователя и программы.

# Список литературы к реферату

1. А. Наследов SPSS 15 профессиональный статистический анализ данных. Питер 2008.
2. Боровиков В. Statistica. Искусство анализа данных на компьютере: Для профессионалов. Спб., 2003.
3. Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка и др. Факторный, дискриминантный и кластерный анализ. М., 1989.
4. CatZone.ws/su – Все новинки интернета бесплатно! [Электрон. ресурс] / Режим доступа: http://catzone.ws/soft/20536-spss-14.0-for-windows-evaluation-version.html
5. Raynald’s SPSS Tools по-русски [Электрон. ресурс] / Режим доступа: http://www.spsstools.ru/FAQ.htm
6. SPSS.ru [Электрон. ресурс] / Режим доступа: http://www.spss.ru/products/spss/add\_on.htm
7. Бесплатная электронная интернет библиотека [Электрон. ресурс] / Режим доступа: http://www.zipsites.ru/psy/psylib/info.php?p=2082
8. Высшая аттестационная комиссия республики Беларусь [Электрон. ресурс] / Режим доступа: http://www.vak.org.by/index.php?go=Pages&in=view&id=54
9. Сообщество HR–менеджеров HR-portal.ru [Электрон. ресурс] / Режим доступа: http://www.hr-portal.ru/spss/Glava1/Index2.php
10. Электронные книги - Computers.plib.ru [Электрон. ресурс] / Режим доступа: http://computers.plib.ru/math/SPSS/Glava%201/Index4.html
11. [Электрон. ресурс] / Режим доступа: http://www.onegid.com/soft/60665-spss-statistics-17.0-mnogojazychnaja-versija.html

# Предметный указатель к реферату

S

SPSS · 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 23, 24, 25

W

Windows · 7, 8 -, - 9 -, - 10 -

Д

Дискриминантный анализ · 17, 18, 19, 20, 21

# Интернет ресурсы в предметной области исследования

1. http://www.google.com.by/

Всемирно известная поисковая система Google. Позволяет производить простой поиск по ключевым словам, при вводе которых мгновенно предлагается перечень вариаций. Очень удобная опция возможность расширенного поиска по группам (среди книг, фото, аудио или видеофайлов, новостей и т.д.), особым признакам (определение, тип файла) в общей сложности более двадцати специализированных сервисов, а также, собственный почтовый сервис. Легкость написания и запоминания выгодно отличает Google от других поисковых систем. Простота оформления и кайзен (постоянное совершенствование во всем) давно покорила сердца и умы миллионов пользователей по всему миру.

2. http://nigma.ru/

Нигма - интеллектуальная поисковая система - это поисковик научно-прикладной характера: она является своего рода исследовательской лабораторией для студентов и аспирантов МГУ. Имеется коммерческая составляющая проекта - реклама от Яндекс.Директа. Главное - это построение эффективного поиска на основе новых поисковых алгоритмов: в частности, это метод, основанный на кластеризации документов. При поисковом запросе пользователя формируется ряд документов, которые группируются в тематические классы, имеющие определенное описание. Таким образом, пользователь может уточнить условия поиска, убрав из него определенные кластеры (сняв галочки перед их названиями). Использует индексную базу нескольких поисковых машин: Google, Yahoo, MSN, Yandex, Rambler, Altavista, Aport. Также имеется и собственная документальная база. Основную аудиторию поисковой системы составляют студенты. Общая популярность Нигмы среди пользователей Рунета остается на достаточно низком уровне.

Особенно приятный момент – «Безопасный Интернет - детям!»

3. http://www.vak.org.by

Сайт Высшей аттестационной комиссии Республики Беларусь. Совершенно необходим каждому ученому! Содержит максимально полный перечень материалов, касающиеся подготовки научных кадров, присуждения ученых степеней и званий. Для начинающих работу представлены краткие паспорта специальностей и программы-минимумы кандидатских экзаменов по специальности. Важный раздел «Каталог файлов» - в нем представлены доступные для скачивания файлы нормативных документов с приложениями и шаблоны регистрационных документов. Организован поиск по сайту и в сети Интернет.

4. http://ru.wikipedia.org/

Википедия. Свободная интернет энциклопедия. Самая большая коллекция знаний человечества, когда-либо собранных вместе: более 14 миллионов энциклопедических статей на 270 языках, с продолжающимся ростом и ежедневным улучшением самими пользователями (специалистами в своей области). Википедия написана понятным языком, и даже самые сложные тематические статьи хорошо адаптированы для понимания. Важным моментом является возможность пояснения каждого термина. Поиск и словарь удобно дополняют друг друга. Особенно приятный момент – «Вики» (так ее прозвали пользователи) не содержит рекламы! Основатель Википедии Джимми Уэйлс.

«Википедия очень важна для всех. Она является достижением человечества» - Fernando Borba и нельзя сказать лучше!

1. http://www.umkc.edu/

Официальный сайт ботанического сада Миссури. Собственная e-net библиотека: редкие книги, фрагменты статей, изображения растений, множество журналов, базы данных.

6. http://herba.msu.ru/russian/index.html

Ботанический сервер московского государственного университета. Содержит перечень ботанических изданий, ботанические подразделения, новости ботанической жизни, ссылки на ботанические ресурсы, библиотеку изображений и прочее. Ресурс доступен на 8-ми языках.

7. http://minpriroda.by/ru

Официальный сайт Министерства природы и охраны окружающей среды Республики Беларусь. На сайте расположена информация о самом Министерстве, об охраняемых и заповедных территориях, памятниках природы, полезные ссылки, справочная информация, красная книга, экологический календарь и экологический бюллетень, актуальные новости, оснащена собственной поисковой системой. Размещена реклама господдержки.

# Действующий личный сайт в WWW

http://veronika-diamond-pinchuk.narod.ru/ - адрес личного сайта.

Здесь вы увидите: мою автобиографию, мои хобби, выпускную работу по ОИТ, презентацию магистерской диссертации, граф научных интересов, гостевую книгу и другую, очень любопытную информацию.

# Презентация магистерской диссертации

Презентацию магистерской диссертации можно посмотреть на сайте http://veronika-diamond-pinchuk.narod.ru/ present.ppt

# Тестовые вопросы по Основам информационных технологий

http://veronika-diamond-pinchuk.narod.ru/test.txt

<question type="close" id="100">

<text>(Пинчук) Как называется органелла, которая осуществляет процесс фотосинтеза?</text>

<answers type="request">

<answer id="1" right="0"><лейкопласт></answer>

<answer id="2" right="0"><хлорофилл></answer>

<answer id="3" right="0"><гемоглобин></answer>

**<answer id="4" right="1"><хлоропласт></answer>**

</answers>

</question>

<question type="close" id="600">

<text> (Пинчук) Для создания заголовка нельзя применить тег?</text>

<answers type="request">

<answer id="302596" right="0">H1</answer>

**<answer id="302597" right="1">H0</answer>**

<answer id="302598" right="0">H10</answer>

<answer id="302599" right="0">H</answer>

</answers>

</question>

# Граф научных интересов

http://veronika-diamond-pinchuk.narod.ru/graph.doc

Граф научных интересов

Магистранта Пинчук В.В. биологический факультет

Специальность ботаника

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Смежные специальности**   |  | | --- | | * **03.00.03 – молекулярная биологи** | | 1. **03.00.03 – молекулярная биология** 2. **03.00.15 – генетика** 3. **03.00.16 – экология** 4. **03.00.28 – биоинформатика** 5. **03.00.32 – биологические ресурсы** | | **Основная специальность**   |  | | --- | | **03.00.05 – Ботаника** | | 1. **систематика высших растений** 2. **популяционная экология** 3. **геоботаника** | | **Сопутствующие специальности**   |  | | --- | | * **01.01.05 – математическая статистика** | | 1. **01.01.05 – теория вероятностей и математическая статистика** 2. **06.01.11 – защита растений** 3. **06.03.02 – лесоустройство и лесная таксация** | |

# Приложение

