

# ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

---

Обзорная статья

УДК 004.428

DOI: 10.18101/2304-5728-2023-3-78-89

## СРАВНИТЕЛЬНЫЙ ОБЗОР СТАТИСТИЧЕСКИХ ПАКЕТОВ ДЛЯ АНАЛИЗА ДАННЫХ

© **Демаков Владимир Иванович**

кандидат технических наук, доцент,

Иркутский государственный медицинский университет Минздрава России

Россия, 664003, г. Иркутск, ул. Красного Восстания, 1

demakovvi@yandex.ru

© **Ларионова Елена Юрьевна**

доктор химических наук, доцент,

Восточно-Сибирский институт МВД России

Россия, 664074, г. Иркутск, ул. Лермонтова, 110

lari555@mail.ru

© **Голодков Юрий Эдуардович**

кандидат технических наук, доцент,

Иркутский национальный исследовательский технический университет

Россия, 664074, г. Иркутск, ул. Лермонтова, 83

utg27@mail.ru

© **Рерке Виктория Игоревна**

кандидат психологических наук, доцент,

Педагогический институт Иркутского государственного университета

Россия, 664003, г. Иркутск, ул. Карла Маркса, 1

rerkew@mail.ru

**Аннотация.** Статья посвящена обзору программных продуктов, предназначенных для анализа данных, использование которых является важным шагом в любом исследовательском проекте, поскольку позволяет исследователю осмыслить собранные данные и сделать содержательные выводы. Большое количество данных, генерируемых каждый день, привело к увеличению спроса на эффективное программное обеспечение для анализа данных. На сегодняшний день доступен широкий спектр статистических пакетов, помогающих в процессе анализа данных, каждый из которых предлагает различные функции и возможности. В этой работе представлены обзор и сравнение некоторых наиболее популярных доступных статистических пакетов и облачных платформ анализа данных. С целью краткого сопоставления возможностей рассматриваемых ресурсов для некоторых пакетов приведены возможности аппроксимации статистических данных рядом Фурье.

Сравнение статистических пакетов основывалось на периоде их создания, стоимости, популярности и предметных областях, в которых они обычно используются. Информация собиралась из различных источников, включая официальные сайты программного обеспечения и статьи в научных журналах.

**Ключевые слова:** статистика, программные комплексы, анализ данных, Big Data, обработка данных.

#### **Для цитирования**

*Демаков В. И., Ларионова Е. Ю., Голодков Ю. Э., Рерке В. И.* Сравнительный обзор статистических пакетов для анализа данных // Вестник Бурятского государственного университета. Математика, информатика. 2023. № 3. С. 78–89.

Сложно с уверенностью выделить статистические программы, используемые специалистами в тех или иных отраслях знаний. Выбор программного обеспечения может варьироваться в зависимости от конкретных особенностей решаемых задач, навыков исследователя, возможностей и ресурсов организации, реализующей проект. Например, юрист, участвующий в сложном коллективном иске, может использовать более продвинутые статистические пакеты, такие как SAS, в то время как юрист в небольшой юридической фирме может использовать Microsoft Excel для подготовки небольших обзоров, содержащих анализ данных. Кроме того, распространенность определенного программного обеспечения может варьироваться в зависимости от региона, поскольку одно программное обеспечение может быть более популярным в одних территориях, чем в других.

Обработка и анализ данных играют решающую роль в различных областях, включая бизнес, социальные науки и науки о жизни. С увеличением доступности данных растет потребность в инструментах и программном обеспечении для эффективной обработки, анализа и визуализации данных. В этой статье мы сравниваем несколько широко используемых пакетов статистической обработки данных и облачных платформ анализа данных. Наша цель — предоставить обзор этих инструментов и выделить их сильные и слабые стороны, стоимость, популярность и области применения.

Microsoft Excel — это программа для работы с электронными таблицами, созданная корпорацией Microsoft в 1985 году. Она широко используется для анализа данных, финансового моделирования и бизнес-аналитики. Excel позволяет пользователям выполнять различные задачи анализа данных, включая манипулирование данными, построение графиков, организация данных, сортировка, фильтрация и формирование сводных таблиц. Программа также поддерживает надстройки и макросы, которые можно использовать для расширения его функциональности. Excel включен в пакет Microsoft Office и доступен как отдельная программа за 139,99 долларов США. Он широко используется во всем мире, что делает его одним из самых популярных инструментов анализа данных. Для эф-

эффективного использования Excel требует базового понимания концепций анализа данных и некоторых навыков в применении его функций и инструментов [1-3]. Также необходимо отметить, что существует большое количество аналогов Excel, многие из которых являются свободно распространяемыми, например Jotform Tables, OpenOffice, Google Sheets, LibreOffice и др. Ввиду большого распространения в России именно табличная структура Excel является стандартом в изучении программных инструментов обработки данных.

Инструмент анализа Фурье в Excel находится в пакете инструментов анализа данных и называется «Анализ Фурье». Чтобы использовать его, достаточно выделить набор данных, нажать кнопку «Анализ данных», выбрать «Анализ Фурье» и заполнить необходимые входные параметры, такие как количество членов в аппроксимации.

SPSS («Statistical Package for the Social Sciences» — статистический пакет для общественных наук) SPSS — пакет статистического программного обеспечения, созданный в 1968 году компанией SPSS Inc. (сегодня принадлежащей IBM). Он широко используется для анализа данных и статистического моделирования в различных областях, включая медицинские науки, социальные науки, образование и маркетинг. SPSS предлагает удобный интерфейс и широкий спектр статистических методов, включая описательную статистику, регрессионный анализ и проверку гипотез. SPSS предоставляется по подписке по цене от 99 долларов за пользователя в месяц. Он широко используется во всем мире, особенно в исследованиях в области социальных наук. Для пользователя требуется более глубокое понимание статистических концепций и методов, а использование этих инструментов требует специальной подготовки и опыта [4-7].

В SPSS вы можете использовать процедуру FFT, перейдя в модуль Спектральный анализ. Чтобы выполнить FFT, вам нужно будет указать набор входных данных, количество терминов в приближении и любые другие соответствующие параметры. Пример кода в SPSS:

```
FFT          VARIABLES=data_set(1)          /OUT=out_data_set(1)
/TERMS=number_of_terms.
```

SAS (система статистического анализа) SAS — пакет статистического программного обеспечения, созданный в 1976 году Институтом SAS. Он широко используется для прогнозной аналитики, управления данными и бизнес-аналитики, особенно в финансах, здравоохранении и розничной торговле. SAS предлагает широкий спектр аналитических методов и поддерживает обработку и визуализацию данных. Он основан на подписке, а цены варьируются в зависимости от продукта и использования. SAS широко используется во всем мире, особенно в США и Европе [8-11].

В SAS вы можете использовать процедуру PROC SPECTRUM для выполнения анализа Фурье. Пример кода в SAS:

```
PROC SPECTRUM DATA=data_set;  
  VAR variable;  
  SPECTRUM variable / N=number_of_terms;  
RUN;
```

RapidMiner — это пакет программного обеспечения для интеллектуального анализа данных, созданный в 2006 году Инго Мирсвой и Тобиасом Куссом. Он широко используется для прогнозной аналитики, машинного обучения и интеллектуального анализа данных. RapidMiner предлагает удобный интерфейс и поддерживает различные форматы данных, включая большие данные. Он основан на подписке, а цены начинаются от 1299 долларов в год. RapidMiner широко используется во всем мире, особенно в Европе и США. Инструментарий RapidMiner предоставляет графический пользовательский интерфейс, который делает анализ данных более доступным для пользователей без специальной подготовки по статистике [12-16].

В RapidMiner вы можете использовать оператор преобразования Фурье для выполнения анализа Фурье. Для этого в RapidMiner нужно:

1. Загрузите свой набор данных в RapidMiner.
2. Перетащите оператор преобразования Фурье на холст процесса.
3. Подключите набор данных к оператору преобразования Фурье.
4. Укажите количество членов в аппроксимации и любые другие соответствующие параметры в окне свойств оператора преобразования Фурье.

Tableau — это программный пакет для визуализации данных и бизнес-аналитики, созданный в 2003 году Кристианом Шаботом, Пэтом Ханраханом и Крисом Столте. Он широко используется для визуализации данных и бизнес-аналитики, особенно в сфере финансов, здравоохранения и розничной торговли. Tableau предлагает широкий спектр вариантов визуализации, включая графики, карты и информационные панели. Он основан на подписке, цены начинаются от 70 долларов за пользователя в месяц. Tableau широко используется во всем мире, особенно в США [17-19].

В Tableau вы можете использовать вычисляемые поля для выполнения анализа Фурье. Пример того, как выполнить анализ Фурье в Tableau:

1. Загрузите свой набор данных в Tableau.
2. Создайте вычисляемое поле, которое выполняет преобразование Фурье в наборе данных.
3. Создайте еще одно вычисляемое поле, указывающее количество терминов в приближении.
4. Используйте эти вычисляемые поля в своих визуализациях.

MINITAB — пакет статистического программного обеспечения, созданный в 1972 году Райаном и Барбарой Джон. Он широко используется для статистического анализа, контроля качества и улучшения процессов,

особенно в производстве и разработке. MINITAB предлагает удобный интерфейс и поддерживает различные статистические методы, включая проверку гипотез и регрессионный анализ. MINITAB доступен для покупки по цене от 895 долларов. Он широко используется во всем мире, особенно в Северной Америке [20-22].

В MINITAB можно использовать процедуру быстрого преобразования Фурье (FFT) для выполнения анализа Фурье. Вот пример кода в MINITAB:

```
FFT data_set, number_of_terms;
```

Stata — пакет статистического программного обеспечения, созданный в 1985 году компанией StataCorp LLC. Он широко используется для анализа данных, статистического моделирования и графики, особенно в экономике, политологии и социологии. Stata предлагает широкий спектр аналитических методов и поддерживает манипулирование и визуализацию данных. Он основан на подписке, а цены начинаются с 45 долларов в месяц. Stata широко используется во всем мире, особенно в академических кругах и исследованиях [10, 23].

В Stata вы можете использовать команду `fft` для выполнения анализа Фурье. Пример кода в Stata:

```
fft variable, nterms(number_of_terms)
```

Gephi — это программный пакет для сетевого анализа и визуализации, созданный в 2008 году Матье Бастианом, Алексисом Жакоми и Гийомом Плик. Он широко используется для сетевого анализа и визуализации, особенно в области социологии, биологии и информатики. Gephi предлагает удобный интерфейс и поддерживает различные методы сетевого анализа, включая обнаружение сообщества и меры центральности. Gephi является бесплатным и открытым исходным кодом. Он широко используется во всем мире, особенно в Европе [24-27].

В Gephi вы можете использовать плагин преобразования Фурье для выполнения анализа Фурье. Пример того, как выполнить анализ Фурье в Gephi:

1. Загрузите свой набор данных в Gephi.
2. Установите плагин преобразования Фурье.
3. Примените плагин преобразования Фурье к набору данных.
4. Укажите количество членов в аппроксимации и любые другие соответствующие параметры в окне свойств плагина преобразования Фурье.

GraphPad — это пакет программного обеспечения для анализа научных данных, созданный в 1984 году компанией GraphPad Software Inc. Он широко используется для анализа научных данных и построения графиков, особенно в науках о жизни и биостатистике. GraphPad предлагает удобный интерфейс и поддерживает различные статистические методы, вклю-

чая t-тесты, ANOVA и регрессионный анализ. GraphPad доступен для покупки по цене от 99 долларов за пользователя. Он широко используется во всем мире, особенно в академических кругах и исследованиях. Использование GraphPad, как и трех описанных выше, требует определенного уровня подготовки и понимания применяемых статистических методов [28-30].

В GraphPad вы можете использовать функцию быстрого преобразования Фурье (FFT) для выполнения анализа Фурье. Пример выполнения анализа Фурье в GraphPad:

1. Загрузите набор данных в GraphPad.
2. Выберите функцию быстрого преобразования Фурье (FFT) в соответствующем меню.
3. Укажите количество членов в аппроксимации и любые другие соответствующие параметры.

Statistica — пакет статистического программного обеспечения, созданный в 1984 году компанией StatSoft (в настоящее время принадлежит Dell). Он широко используется для анализа данных, прогнозного моделирования и интеллектуального анализа данных. Statistica предлагает широкий спектр статистических методов и поддерживает визуализацию данных. Он доступен для покупки по цене от 1895 долларов. Statistica широко используется во всем мире, особенно в научных кругах и исследованиях [31-33].

В Statistica вы можете использовать процедуру преобразования Фурье для выполнения анализа Фурье. Вот пример кода в Statistica:

```
FourierTransform data_set, number_of_terms;
```

R — это бесплатный язык программирования с открытым исходным кодом для статистических вычислений и графики, созданный в 1993 году Россом Ихакой и Робертом Джентльменом. Он широко используется для анализа данных, статистического моделирования и машинного обучения, особенно в научных кругах и исследованиях. R предлагает широкий спектр статистических методов и поддерживает визуализацию данных. Он широко используется во всем мире, особенно в академических кругах и исследованиях. Для эффективного применения этого ресурса пользователю нужно владеть минимальными навыками программирования и пониманием основ статистического анализа [34].

В R вы можете использовать функцию `fft` из пакета `stats` для выполнения анализа Фурье. Пример кода на R:

```
data_set <- read.csv("data_set.csv")  
Fourier_transform <- stats::fft(data_set, number_of_terms)
```

Microsoft Azure — это платформа облачных вычислений, созданная Microsoft в 2010 году. Она широко используется для анализа данных и облачных вычислений, особенно в сферах бизнеса, финансов и здравоохранения. Microsoft Azure предлагает различные услуги для анализа данных, включая хранение данных, машинное обучение и визуализацию дан-

ных. Microsoft Azure предоставляется по подписке, и цены варьируются в зависимости от использования. Он широко используется во всем мире, особенно в США и Европе [35-38].

В Microsoft Azure, как и в Amazon Web Services, вы можете использовать службу машинного обучения Azure для выполнения анализа Фурье. Пример выполнения анализа Фурье в Azure:

1. Загрузите набор данных в Azure.
2. Создайте рабочую область машинного обучения Azure.
3. Создайте новый эксперимент в рабочей области машинного обучения Azure.
4. Перетащите модуль преобразования Фурье на холст эксперимента.
5. Подключите набор данных к модулю преобразования Фурье.
6. Укажите количество членов в аппроксимации и любые другие соответствующие параметры в окне свойств модуля преобразования Фурье.

Веб-сервисы Amazon Web Services (AWS) — это платформа облачных вычислений, созданная Amazon в 2002 году. Она широко используется для анализа данных и облачных вычислений, особенно в сферах бизнеса, финансов и здравоохранения. AWS предлагает различные сервисы для анализа данных, включая хранение данных, машинное обучение и визуализацию данных. AWS предоставляется по подписке, и цены варьируются в зависимости от использования. Он широко используется во всем мире, особенно в США и Европе. Облачные технологии требуют обучения работе с конкретными инструментами и услугами, предлагаемыми каждой платформой [36].

Наряду с широко распространёнными статистическими пакетами, позволяющими решать задачи широкого профиля, в практике отдельных ведомств используются специализированные информационные системы. Так, например, российские юридические органы внедрили автоматизированную информационную систему, позволяющую накапливать данные правовой природы, а также предоставлять авторизованным пользователям богатый инструментарий для обработки и визуализации статистики. Государственная автоматизированная система правовой статистики (ГАС ПС) в России представляет собой централизованную базу данных, которая собирает и хранит статистические данные, касающиеся правовой системы и отправления правосудия в стране. ГАС ПС используется для сбора, обработки и анализа статистических данных из различных источников, включая суды, прокуратуры и другие юридические организации. Данные, собираемые через ГАС ПС, используются для мониторинга деятельности правовой системы, оценки эффективности юридических процедур и поддержки процессов принятия решений, связанных с отправлением правосудия [39-41].

ГАС ПС спроектирован так, чтобы быть удобным для пользователя, с графическим пользовательским интерфейсом, который позволяет пользователям легко получать доступ к данным и выполнять различные задачи анализа. Программный ресурс ГАС ПС поддерживает различные статистические методы и методы, включая описательную статистику, логическую статистику и прогнозную аналитику. Программное обеспечение также предоставляет различные инструменты визуализации, такие как диаграммы и графики, которые помогают пользователям анализировать и интерпретировать данные. Чтобы эффективно использовать программные ресурсы ГАС ПС, пользователи должны иметь базовое понимание статистических концепций и методов, а также некоторую подготовку по использованию программного обеспечения. ГАС ПС обеспечивает обучение и поддержку своих пользователей, а технические специалисты готовы помочь пользователям с любыми вопросами или проблемами, которые у них могут возникнуть [41].

В целом ГАС ПС является важным инструментом федерального уровня. Использование программного ресурса ГАС ПС позволяет пользователям централизованно и эффективно получать доступ к данным, связанным с правовой системой, и анализировать их, помогая повысить эффективность юридических процедур и поддерживать процессы принятия решений [42]. Обучающиеся в образовательных организациях МВД России уже в период обучения знакомятся с возможностями ГАС ПС.

Аналогичные информационные ресурсы создаются в других отраслях. Например, медицинский информационно-аналитический центр (МИАЦ) в Министерстве здравоохранения.

В заключении отметим, что существует широкий спектр доступного программного обеспечения для анализа данных, каждое из которых имеет свои особенности, возможности и уровень подготовки, необходимый для их эффективного использования [43]. Microsoft Excel — популярный выбор для простых задач анализа данных, в то время как SPSS, SAS и Statistica — это специализированное программное обеспечение, требующее специальной подготовки и опыта работы со статистическими методами. RapidMiner, Tableau и MINITAB — это удобное программное обеспечение для анализа данных, а Stata, Gephi, GraphPad и R предназначены для конкретных областей применения. Microsoft Azure и Amazon Web Services — это облачные платформы для анализа данных, которые предоставляют мощные вычислительные ресурсы и варианты хранения данных. Каждый из этих инструментов имеет свои сильные и слабые стороны и используется в различных предметных областях. Важно отметить, что ни один пакет не является лучшим для всех целей, и выбор пакета будет зависеть от конкретных требований исследовательского проекта. Лучший инструмент для конкретной задачи зависит от конкретных потребностей и требований пользователя [44]. Немаловажным также является и возможность использования рассматриваемых ресурсов для образовательных це-



лей [45]. Стоимость также является фактором, который следует учитывать, поскольку некоторые инструменты доступны по подписке, а другие доступны для покупки. В целом, выбор инструмента анализа данных должен основываться на конкретных потребностях и целях пользователя и данных, с которыми он работает.

Необходимо также отметить, что приведенный обзор содержит лишь часть широко применяемых на практике статистических пакетов. Можно указать множество распространенных и мощных ресурсов анализа и обработки данных, например программы Mathcad, Matlab, Maple и многие другие. Кроме этого, различные государственные структуры используют специализированные инструменты для работы с большими массивами данных.

#### Литература

1. Microsoft Excel official website. URL: <https://products.office.com/en-us/excel> (дата обращения: 06.02.2023).
2. Microsoft Excel tutorial for legal research. URL: <https://www.gcflearnfree.org/excel-for-legal-professionals/> (дата обращения: 06.02.2023).
3. Вадзинский Р. Н. Статистические вычисления в среде Excel: Библиотека пользователя. Санкт-Петербург: Питер, 2010. 608 с.
4. SPSS official website. URL: <https://www.ibm.com/analytics/spss-statistics-software> (дата обращения: 06.02.2023).
5. Наследов А. SPSS 19: профессиональный статистический анализ данных. Санкт-Петербург: Питер, 2011. 400 с.
6. Бююль А., Цефель П. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. Санкт-Петербург: Диа Софт ЮП, 2005. 608 с.
7. Field A. Discovering Statistics Using SPSS for Windows: Advanced Techniques for Beginners. Sage Publications, 2000. 512 p.
8. SAS official website. URL: [https://www.sas.com/en\\_us/home.html](https://www.sas.com/en_us/home.html) (дата обращения: 06.02.2023).
9. SAS tutorials for legal research. URL: <https://support.sas.com/en/education/courses/legal-analytics.html> (дата обращения: 06.02.2023).
10. Унгуриян Т. Н., Гржибовский А. М. Корреляционный анализ с использованием пакета статистических программ STATA // Экология человека. 2014. № 9. С. 60–64.
11. Susan J. and Lora D. Delwiche. The Little SAS Enterprise Guide Book // SAS Institute Inc., Cary, NC, USA, 2017. P. 146–152.
12. Rapidminer. URL: <https://my.rapidminer.com/nexus/account/index.html#downloads> (дата обращения: 06.02.2023).
13. RapidMiner tutorials for legal research. URL: <https://docs.rapidminer.com/studio/tutorials/> (дата обращения: 06.02.2023).
14. Интеллектуальный анализ данных с помощью пакета RapidMiner : учебно-методическое пособие / А. Г. Степанов, В. С. Блюм, В. С. Васильева [и др.]. Санкт-Петербург: Изд-во Санкт-Петерб. гос. ун-та аэрокосмического приборостроения, 2021. 251 с.

15. Крахалёв В., Введение в RapidMiner. URL: <https://habrahabr.ru/post/269427/> (дата обращения: 06.02.2023).
16. Hofmann M., Klinkenberg R., RapidMiner: Data Mining Use Cases and Business Analytics Applications. 1st Edition / Chapman & Hall/CRC, 2013. 525 p.
17. Tableau official website. URL: <https://www.tableau.com/> (дата обращения: 06.02.2023).
18. Cintas P. G. Industrial Statistics with Minitab. Chichester, West Sussex: Wiley, 2012. 420 p.
19. Анализ данных в Tableau на практике. 100 советов, уроков и стратегий от мастера дзен в Tableau / пер. с англ. А. Ю. Гинько. Москва: ДМК Пресс, 2021. 546 с.
20. Meyer Ruth K. Minitab Guide to Statistics / Ruth K. Meyer, David D. Krueger. 3rd. Upper Saddle River, NJ: Prentice-Hall Publishing. 2004. 448 p.
21. MINITAB official website. URL: <https://www.minitab.com/> (дата обращения: 06.02.2023).
22. Беликов В. В., Листопад А. А., Милютин Е. М. Обзор статистического пакета Minitab // Инновационные направления разработки и использования информационных технологий: материалы II Международной заочной студенческой научно-практической конференции (Брянск, 23-25 мая 2015 г.). Брянск: Изд-во Брянского ГАУ, 2016. С. 178–180.
23. Аль Е. Х. А. А. Comparison of statistical functions for programs (SAS, SPSS, and MINITAB) // Молодой ученый. 2013. № 5. С. 131–142.
24. Gephi official website. URL: <https://gephi.org/> (дата обращения: 06.02.2023).
25. Питиляк Д. А., Рожкова А. О. Средства визуализации данных Gephi и Google в экономических исследованиях // Молодой ученый. 2016. № 12(116). С. 1408–1412.
26. Gephi как средство визуализации данных. URL: <https://habr.com/post/136575/> (дата обращения: 06.02.2023).
27. Филяк П. Ю., Тебеньков Н. В., Королев С. В. Графовая среда Gephi в обеспечении информационной безопасности // Информация и безопасность. 2019. Т. 22, № 1. С. 102–107.
28. GraphPad official website. URL: <https://www.graphpad.com/> (дата обращения: 06.02.2023).
29. GraphPad tutorials for legal research. URL: <https://www.graphpad.com/support/tutorials> (дата обращения: 06.02.2023).
30. Motulsky H. J., Christopoulos A. Fitting Models to Biological Data Using Linear and Nonlinear Regression. A Practical Guide to Curve Fitting. San Diego CA: GraphPad Software Inc, 2003.
31. Statistics textbooks: "Introduction to Statistics" by Robert N. Baumeister and John D. Berry, "An Introduction to Statistical Learning" by Gareth James et al., and "Statistics for People Who (Think They) Hate Statistics" by Neil J. Salkind.
32. Чурилова Э. Ю., Салин В. Н. Статистический анализ данных цифровой экономики в системе "STATISTICA". Москва: Компания КноРус, 2019. 238 с.
33. Боровиков В. STATISTICA. Искусство анализа данных на компьютере: для профессионалов. Санкт-Петербург: Питер, 2003. 688 с.
34. Груздев А. В. Прогнозное моделирование в IBM SPSS Statistics, R и Python. Метод деревьев решений и случайный лес. Москва: ДМК-Пресс, 2018. 642 с.

35. Microsoft Azure official website. URL: <https://azure.microsoft.com/> (дата обращения: 06.02.2023).

36. Open tack Cloud Computing Cookbook: Over 100 practical recipes to help you build and operate Open Stack cloud computing, storage, networking, and automation // Advanced Materials. 2020. Pp. 430.

37. Таллоч Митч. Знакомство с Windows Azure. Для ИТ-специалистов: пер. с англ. Москва: ЭКОМ Паблишерз, 2014. 154 с.

38. Сенчилов В. В., Григорьева Г. М., Ходченков В. Ю. Разработка программного обеспечения, основанная на взаимодействии с облачными сервисами Microsoft Azure // Лучшая научная статья 2017: сборник статей IX Международного научно-практического конкурса (Пенза, 30 мая 2017 г.). Пенза: Наука и просвещение, 2017. С. 28–31.

39. GAS PS official website. URL: <https://pravo.gov.ru/> (дата обращения: 06.02.2023).

40. «Об утверждении Концепции цифровой трансформации органов и организаций прокуратуры до 2025 года»: приказ Генеральной прокуратуры РФ от 14 сентября 2017 г. № 627 (ред. от 01.12.2021). URL: [https://www.consultant.ru/document/cons\\_doc\\_LAW\\_278651](https://www.consultant.ru/document/cons_doc_LAW_278651) (дата обращения: 06.02.2023).

41. Шарипова Д. М. Система ГАС ПС как значимая часть цифровизации органов прокуратуры Шарипова // Вестник науки. 2022. Т. 5, № 5(50). С. 200–203.

42. Бударин И. С. Государственная автоматизированная система правовой статистики в информационном пространстве // Вестник науки и образования. Сер. Право. 2019. № 19 (73). С. 78–80. URL: <https://cyberleninka.ru/article/n/gosudarstvennaya-avtomatizirovannaya-sistema-pravovoy-statistiki-v-informatsionnom-prostranstve/viewer> (дата обращения: 06.02.2023).

43. Черткова Е. А. Статистика. Автоматизация обработки информации: учебное пособие для вузов / под общей редакцией Е. А. Чертковой. 2-е изд., испр. и доп. Москва: Юрайт, 2017. 195 с.

44. Демаков В. И., Ланг Е. П. О роли автоматизации медицинской статистики // Здоровье населения и среда обитания. 2019. № 6(315). С. 4–7.

45. Дисциплина «Правовая статистика» как средство подготовки выпускников вузов МВД России к использованию математических методов в информационно-аналитической работе / Е. Ю. Ларионова, Ю. Э. Голодков, В. И. Демаков, Я. А. Портная // Вестник Красноярского государственного педагогического университета им. В. П. Астафьева. 2021. № 1(55). С. 36–45. DOI: 10.25146/1995-0861-2021-55-1-257.

*Статья поступила в редакцию 28.06.2023; одобрена после рецензирования 07.07.2023; принята к публикации 27.09.2023.*

#### COMPARATIVE REVIEW OF STATISTICAL PACKAGES FOR DATA ANALYSIS

*Vladimir I. Demakov*

Candidate of Technical Sciences, Associate Professor,  
Irkutsk State Medical University of the Ministry of Health of Russia  
1 Krasnogo Vosstaniya str., Irkutsk 664003, Russia  
[demakovvi@yandex.ru](mailto:demakovvi@yandex.ru)

*Elena Yu. Larionova*

Doctor of Chemical Sciences, Associate Professor,  
East-Siberian Institute of the Ministry of Internal Affairs of Russia  
110 Lermontov str., Irkutsk 664074, Russia  
lari555@mail.ru

*Yuri E. Golodkov*

Candidate of Technical Sciences, Associate Professor,  
Irkutsk National Research Technical University  
83 Lermontov str., Irkutsk 664074, Russia  
yrg27@mail.ru

*Victoria I. Rerke*

Candidate of Psychological Sciences, Associate Professor,  
Pedagogical Institute of Irkutsk State University  
1 Karl Marx str., Irkutsk, Russia, 664003,  
rerkew@mail.ru

*Abstract.* The article is devoted to an overview of software products designed for data analysis, the use of which is an important step in any research project, since it allows the researcher to comprehend the collected data and draw meaningful conclusions. The large amount of data generated every day has led to an increased demand for efficient data analysis software. Today, a wide range of statistical packages are available to help in the process of data analysis, each of which offers different functions and capabilities. This paper provides an overview and comparison of some of the most popular available statistical packages and cloud-based data analysis platforms. In order to briefly compare the capabilities of the resources under consideration, for some packages, the possibilities of approximating statistical data by a Fourier series are given.

Comparison of statistical packages was based on the period of their creation, cost, popularity and subject areas in which.

*Keywords:* statistics, software packages, data analysis, Big Data, data processing.

*For citation*

*Demakov V. I., Larionova E. Yu., Golodkov Yu. E., Rerke V. I. Comparative Review of Statistical Packages for Data Analysis // Bulletin of Buryat State University. Mathematics, Informatics. 2023. N. 3. P. 78–89.*

*The article was submitted 28.06.2023; approved after reviewing 07.07.2023; accepted for publication 27.09.2023.*