ФЕДЕРАЛЬНОЕ АГЕНТСТВО ЖЕЛЕЗНОДОРОЖНОГО ТРАНСПОРТА

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Петербургский государственный университет путей сообщения

Императора Александра I»

(ФГБОУ ВО ПГУПС)

Кафедра «Математика и моделирование»

**РАБОЧАЯ ПРОГРАММА**

*дисциплины*

«АНАЛИЗ ДАННЫХ» (Б1.Б.18)

для направления

38.03.05 «Бизнес-информатика»

по профилю

«Архитектура предприятия»

Форма обучения – очная

Санкт-Петербург

2018



ЛИСТ СОГЛАСОВАНИЙ

Рабочая программа рассмотрена, обсуждена на заседании кафедры

«Математика и моделирование»

Протокол № \_6\_ от «\_19\_» \_января\_ 2018 г.

|  |  |  |
| --- | --- | --- |
| Заведующий кафедрой«Математика и моделирование» | \_\_\_\_\_\_\_\_\_\_\_ | В.А. Ходаковский |
| «\_19\_» \_января\_ 2018 г. |  |  |

|  |  |  |
| --- | --- | --- |
| СОГЛАСОВАНО |  |  |
| Председатель методической комиссии факультета «Промышленное и гражданское строительство» | \_\_\_\_\_\_\_\_\_\_\_\_ | Р.С. Кударов |
| «\_19\_» \_января\_ 2018 г. |  |  |
|  |  |  |
| Руководитель ОПОП | \_\_\_\_\_\_\_\_\_\_ | В.А. Ходаковский |
| «\_19\_» \_января\_ 2018 г. |  |  |

**1. Цели и задачи дисциплины**

Рабочая программа составлена в соответствии с ФГОС ВО, утвержденным «11» августа 2016 г., приказ № 1002 по направлению38.03.05 «Бизнес-информатика», по дисциплине «Анализ данных».

Целью изучения дисциплины является обучение студентов основным моделям и методам сбора, обработки и анализа реальных социально-экономических данных, формирование у студентов умений в данной области, необходимых для профессиональной деятельности по направлению «Бизнес-информатика».

Для достижения поставленной цели решаются следующие задачи:

* освоение студентами вероятностного подхода к анализу реальных социально-экономических явлений;
* формирование у студентов системного подхода к применению методов анализа данных;
* овладение методами и алгоритмами обработки и анализа данных, содержащих случайные погрешности;
* обучение основам статистического анализа данных с элементами многомерного анализа;
* обучение студентов навыкам анализа данных с использованием специализированных систем анализа данных (R, MatLAB, Statistica, SPSS).

**2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения основной профессиональной образовательной программы**

Планируемыми результатами обучения по дисциплине являются: приобретение знаний, умений, навыков.

В результате освоения дисциплины обучающийся должен:

**ЗНАТЬ**:

* математические инструментальные средства для обработки, анализа и систематизации информации по теме исследования;
* принципы и методы сбора, систематизации и анализа данных наблюдений;
* основные методы статистического анализа данных;
* основные алгоритмы анализа данных.

**УМЕТЬ**:

* математические инструментальные средства для обработки, анализа и систематизации информации по теме исследования;
* принципы и методы сбора, систематизации и анализа данных наблюдений;
* основные методы статистического анализа данных;
* основные алгоритмы анализа данных.

**ВЛАДЕТЬ**:

* методами проведения факторного и кластерного анализа;
* методами и алгоритмами теории анализа данных;
* техникой анализа данных с использованием специализированных пакетов (R, MatLAB, Statistica, SPSS).

Приобретенные знания, умения, навыки, характеризующие формирование компетенций, осваиваемые в данной дисциплине, позволяют решать профессиональные задачи, приведенные в соответствующем перечне по видам профессиональной деятельности в п. 2.4 общей характеристики основной профессиональной образовательной программы (ОПОП).

Изучение дисциплины направлено на формирование следующих **общепрофессиональных компетенций (ОПК)**:

* способностью решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности (ОПК-1);
* способностью работать с компьютером как средством управления информацией, работать с информацией из различных источников, в том числе в глобальных компьютерных сетях (ОПК-3).

Изучение дисциплины направлено на формирование следующих **профессиональных компетенций (ПК)**, соответствующих виду профессиональной деятельности, на который ориентирована программа бакалавриата:

*Организационно-управленческаядеятельность*:

* проведение обследования деятельности и ИТ-инфраструктуры предприятий (ПК-5);
* управление контентом предприятия и Интернет-ресурсов, процессами создания и использования информационных сервисов (контент-сервисов) (ПК-6);
* использование современных стандартов и методик, разработка регламентов для организации управления процессами жизненного цикла ИТ-инфраструктуры предприятий (ПК-7);
* организация взаимодействия с клиентами и партнерами в процессе решения задач управления жизненным циклом ИТ-инфраструктуры предприятия (ПК-8).

Область профессиональной деятельности обучающихся, освоивших данную дисциплину, приведена в п. 2.1 общей характеристики ОПОП.

Объекты профессиональной деятельности обучающихся, освоивших данную дисциплину, приведены в п. 2.2 общей характеристики ОПОП.

**3. Место дисциплины в структуре основной профессиональной образовательной программы**

Дисциплина «Анализ данных» (Б1.Б.18) относится к базовой части и является обязательной дисциплиной обучающегося.

**4. Объем дисциплины и виды учебной работы**

|  |  |  |
| --- | --- | --- |
| **Вид учебной работы** | **Всего часов** | **Семестр** |
| **5** |
| Контактная работа (по видам учебных занятий)В том числе:* лекции (Л)
* практические занятия (ПЗ)
* лабораторные работы (ЛР)
 | 3216-16 | 3216-16 |
| Самостоятельная работа (СРС) (всего) | 31 | 31 |
| Контроль | 9 | 9 |
| Форма контроля знаний |  | З |
| Общая трудоемкость: час / з.е. | 72/2 | 72/2 |

*Примечания: «Форма контроля знаний» – зачет (З)*

**5. Содержание и структура дисциплины**

5.1 Содержание дисциплины

|  |  |  |
| --- | --- | --- |
| **№ п/п** | **Наименование раздела дисциплины** | **Содержание раздела** |
| 1 | Основные методы статистического анализа данных. | Статистические оценки параметров, их свойства (повторение). Основные методы точечного и интервального оценивания параметров Основы проверки статистических гипотез.  |
| 2 | Методы исследования однородности и независимости двух выборок | Понятие об однородности выборок.Выявление неоднородности, связанной со сдвигом (классический критерий Стьюдента, критерий знаков, критерий Вилкоксона) или масштабом (F- критерий). Проверка однородности против альтернатив общего вида (критерий Колмогорова – Смирнова, однородности χ2). Проверка независимости признаков (критерий независимости χ2) |
| 3 | Анализ главных компонент, корреляционный анализ | Постановка задачи, свойства главных компонент, геометрия главных компонент |
| 4 | Элементы факторного анализа | Модель факторного анализа, неоднозначность решения, уравнения правдоподобия, оценка факторных нагрузок, выбор числа факторов, оценка факторов.  |
| 5 | Элементы дисперсионного анализа | Модель однофакторного анализа, таблицы однофакторного анализа, классический F-критерий, сбалансированные планы. Модель двухфакторного дисперсионного анализа, оценка параметров, проверка гипотез, F-критерии. |
| 6 | Задачи классификации, алгоритмы кластерного анализа | Эвристические алгоритмы классификации, иерархические алгоритмы, алгоритм k-средних. Критерии качества классификации, выбор числа классов.  |

5.2 Разделы дисциплины и виды занятий

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **№ п/п** | **Наименование раздела дисциплины** | **Л** | **ПЗ** | **ЛР** | **СРС** |
| 1 | Основные методы статистического анализа данных. | 2 | - | 2 | 4 |
| 2 | Методы исследования однородности и независимости двух выборок | 4 | - | 4 | 8 |
| 3 | Анализ главных компонент, корреляционный анализ | 4 | - | 4 | 8 |
| 4 | Элементы факторного анализа | 2 | - | 2 | 3 |
| 5 | Элементы дисперсионного анализа | 2 | - | 2 | 4 |
| 6 | Задачи классификации, алгоритмы кластерного анализа | 2 | - | 2 | 4 |
| **Итого** | **16** | **-** | **16** | **31** |

**6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине**

|  |  |  |
| --- | --- | --- |
| **№****п/п** | **Наименование раздела дисциплины** | **Перечень учебно-методического обеспечения** |
| 1 | Основные методы статистического анализа данных. | 1) Бородин А.Н. Элементарный курс теории вероятностей и математической статистики, 8-е изд., стер.- СПб: Лань, 2011,. – 224 c.2) Горяинова Е.Р., А.Р. Панков, Е.Н. Платонов. Прикладные методы анализа статистических данных. Учебное пособие [Электронный ресурс], . – М.: Издательский дом Высшей школы экономики. 2012, 312 с. http://www.knigafund.ru/books/1724552) Гмурман В.Е. Теория вероятностей и математическая статистика: учеб.пособие/В.Е.Гмурман, - 12-е изд.перераб.-М: Юрайт, 2010.- 479 с. ил. –(Основы наук) Экземпляры: всего 101 – ЧЗОД(1), ОУЛ(100) ) (Шифр: В17 я73 Г55)3)Математическая статистика в системе MatLAB. Лабораторный практикум / сост. Н.В.Грибкова. Методические указания, СПб, ПГУПС, 2013. . – 47 c.  |
| 2 | Методы исследования однородности и независимости двух выборок | 1)Ивченко Г.И., Медведев Ю.И. Введение в математическую статистику, М: ЛКИ, 2010 [Электронный ресурс] <http://www.twirpx.com/file/529807/>Кельберт М.Я., Сухов Ю.И. [Вероятность и статистика в примерах и задачах. Том 1: Основные понятия теории вероятностей и математической статистики](http://biblio.mccme.ru/node/2258) М.: МЦМНО, 2007 . – 456 c.2)Лагутин М.Б. Наглядная математическая статистика. Учеб.пособие/ М.Б.Лагутин.- М.: БИНОМ, Лаборатория знаний, 2007. – 472 с. ил. ISBN 978-5-94774-345-6. [Электронный ресурс] <http://www.1variant.ru/content/uchebniki/matematika/650.pdf> |
| 3 | Анализ главных компонент, корреляционный анализ | 1)Ивченко Г.И., Медведев Ю.И. Введение в математическую статистику, М: ЛКИ, 2010 [Электронный ресурс] <http://www.twirpx.com/file/529807/> |
| 4 | Элементы факторного анализа | 1)Ивченко Г.И., Медведев Ю.И. Введение в математическую статистику, М: ЛКИ, 2010 [Электронный ресурс] <http://www.twirpx.com/file/529807/> |
| 5 | Элементы дисперсионного анализа | 1)Ивченко Г.И., Медведев Ю.И. Введение в математическую статистику, М: ЛКИ, 2010 [Электронный ресурс] <http://www.twirpx.com/file/529807/> |
| 6 | Задачи классификации, алгоритмы кластерного анализа | 1)Лагутин М.Б. Наглядная математическая статистика. Учеб.пособие/ М.Б.Лагутин.- М.: БИНОМ, Лаборатория знаний, 2007. – 472 с. ил. ISBN 978-5-94774-345-6. [Электронный ресурс] <http://www.1variant.ru/content/uchebniki/matematika/650.pdf> |

**7. Фонд оценочных средств для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине**

Фонд оценочных средств по дисциплине является неотъемлемой частью рабочей программы и представлен отдельным документом, рассмотренным на заседании кафедры и утвержденным заведующим кафедрой.

**8. Перечень основной и дополнительной учебной литературы, нормативно-правовой документации и других изданий, необходимых для освоения дисциплины**

8.1 Перечень основной учебной литературы, необходимой для освоения дисциплины

1. Бородин А.Н. Элементарный курс теории вероятностей и математической статистики, 8-е изд., стер.- СПб: Лань, 2011,. – 224 c/
2. Буреева Н.Н. Многомерный статистический анализ с использованием ППП “STATISTICA”. Учебно-методический материал [Электронный ресурс],ННГУ им. Н.И. Лобачевского, Нижний Новгород, 2007. — 112с. http://www.unn.ru/pages/issues/aids/2007/57.pdf
3. [Введение в систему R статистического анализа данных](http://www.google.ru/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDAQFjAA&url=http%3A%2F%2Fmpoctok.narod.ru%2Fr%2Fintro.htm&ei=a88gUfykNaSN4gTikICIDw&usg=AFQjCNHRpnjSGD3eyr8SGS4vv3BJ2SMcfA&sig2=-JdcKaUsjyPcZCVSiCmnbA&bvm=bv.42553238,d.bGE&cad=rjt) (электронный ресурс) <http://mpoctok.narod.ru/r/intro.htm>
4. Горяинова Е.Р., А.Р. Панков, Е.Н. Платонов. Прикладные методы анализа статистических данных. Учебное пособие [Электронный ресурс],. – М.: Издательский дом Высшей школы экономики. 2012, 312 с.http://www.knigafund.ru/books/172455
5. Гмурман В.Е. Теория вероятностей и математическая статистика: учеб.пособие/В.Е.Гмурман, - 12-е изд.перераб.-М: Юрайт, 2010.- 479 с. ил. –(Основы наук) Экземпляры: всего 101 – ЧЗОД(1), ОУЛ(100) ) (Шифр: В17 я73 Г55)
6. Ивченко Г.И., Медведев Ю.И. Введение в математическую статистику, М: ЛКИ, 2010.-600 с. [Электронный ресурс]<http://www.twirpx.com/file/529807/>
7. Фаддеев М.А. Элементарная обработка результатов эксперимента: учеб.пособие/М.А.Фаддеев.- СПб; М.; Краснодар: Лань, 2008. – 117 с. ил. (Учебник для вузов. Специальная литература). Экземпляры: всего 21.- НБ(2), ОУЛ(19). (Шифр: Т43698 Ф15)

8.2 Перечень дополнительной учебной литературы, необходимой для освоения дисциплины

1. Кухаренко Л.А., Малошевский С.Г. Математическая статистика./ учебное пособие.- СПб ПГУПС, 1993. – 142 c.
2. Лагутин М.Б. Наглядная математическая статистика. Учеб.пособие/ М.Б.Лагутин.- М.: БИНОМ, Лаборатория знаний, 2007. – 472 с. ил. ISBN 978-5-94774-345-6. [Электронный ресурс] <http://www.1variant.ru/content/uchebniki/matematika/650.pdf>
3. Письменный Д.Т. Конспект лекций по теории вероятностей и математической статистике и случайным процессам/Д.Т.Письменный.-4-е изд.испр.—Айрис Пресс. 2008. – 287 с. ил. –(Высшее образование) Экземпляры всего – 84 ю- НБ(2), ОУЛ(82) (Шифр: В17 я7-2 П35 Т43651)
4. Калинина В.Н., Соловьев В.И. Введение в многомерный статистический анализ. Учебноепособие/ ГУУ. – М., 2003. – 66с. [Электронный ресурс] <http://visoloviev.ru/booksmath/MSA.pdf>.
5. Семеничин Е.А. Теория вероятностей в примерах и задачах: учеб. Пособие/ Е.А. Семеничин. – СПб; М. ; Краснодар; Лань, 2007. – 351 с.: ил. (Учебник для вузов. Специальная литература) Экземпляры: всего 31 – НБ(1) ОУЛ (30) Гриф – Учебно-методическое обеспечение университетов РФ по прикладной математике (Шифр: В171 я7-4 С30).

8.3 Перечень нормативно-правовой документации, необходимой для освоения дисциплины

Нормативно-правовая документация при освоении дисциплины не используется.

8.4 Другие издания, необходимые для освоения дисциплины

1. Боровиков В.П. Statistica: Искусство анализа данных на компьютере//В.П.Боровиков, СПб.: Питер, 2001, 656 с.
2. Дьяконов В.П. MatLAB 6: Учебный курс. — СПб: Питер, 2001. – 592 c.
3. Кельберт М.Я., Сухов Ю.И. [Вероятность и статистика в примерах и задачах. Том 1: Основные понятия теории вероятностей и математической статистики](http://biblio.mccme.ru/node/2258) М.: МЦМНО, 2007 . – 456 c.
4. Математическая статистика в системе MatLAB. Лабораторный практикум / сост. Н.В.Грибкова. Методические указания, СПб, ПГУПС, 2013. . – 47 c.
5. Модели математической статистики /Сост.Баушев А.Н., Грибкова Н.В., Ингстер Ю.И., Методические указания, СПб, ПГУПС, 2000. – 86 c.
6. Тюрин, Ю.Н. Анализ данных на компьютере / Ю.Н. Тюрин, А.А. Макаров ; под ред. В.Э. Фигурнова. – М. : ИНФРА-М, 2003. – 544 с.

**9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины**

1. Личный кабинет обучающегося и электронная информационно-образовательная среда. [Электронный ресурс]. – Режим доступа: <http://sdo.pgups.ru/> (для доступа к полнотекстовым документам требуется авторизация).
2. Электронно-библиотечная система издательства «Лань». Режим доступа: <http://e.lanbook.com> (для доступа к полнотекстовым документам требуется авторизация).
3. Электронная библиотека ЮРАЙТ. Режим доступа: <https://biblio-online.ru/> (для доступа к полнотекстовым документам требуется авторизация).
4. Электронно-библиотечная система ibooks.ru («Айбукс»). Режим доступа: <https://ibooks.ru/home.php?routine=bookshelf> (для доступа к полнотекстовым документам требуется авторизация).
5. Электронная библиотека «Единое окно доступа к образовательным ресурсам». Режим доступа: <http://window.edu.ru>. – свободный.
6. Электронные словари и энциклопедии на Академике. Режим доступа <https://dic.academic.ru/>. – свободный.

**10. Методические указания для обучающихся по освоению дисциплины**

Порядок изучения дисциплины следующий:

1. Освоение разделов дисциплины производится в порядке, приведенном в разделе 5 «Содержание и структура дисциплины». Обучающийся должен освоить все разделы дисциплины с помощью учебно-методического обеспечения, приведенного в разделах 6, 8 и 9 рабочей программы.
2. Для формирования компетенций обучающийся должен представить выполненные типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков, предусмотренные текущим контролем (см. фонд оценочных средств по дисциплине).
3. По итогам текущего контроля по дисциплине, обучающийся должен пройти промежуточную аттестацию (см. фонд оценочных средств по дисциплине).

**11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем**

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине:

* электронная информационно-образовательная среда Петербургского государственного университета путей сообщения Императора Александра I [Электронный ресурс]. Режим доступа: <http://sdo.pgups.ru>.
* Интернет-сервисы и электронные ресурсы (поисковые системы, электронная почта, онлайн-энциклопедии и словари, электронные учебные и учебно-методические материалы согласно п. 9 рабочей программы).
* ежегодно обновляемый комплект лицензионного программного обеспечения, установленного на технических средствах, размещенных в специальных помещениях и помещениях для самостоятельной работы: операционная система Windows; MS Office; Антивирус Касперский.

**12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине**

Материально-техническая база обеспечивает проведение всех видов учебных занятий, предусмотренных учебным планом по направлению 38.03.05 и соответствует действующим санитарным и противопожарным нормам и правилам.

Она содержит специальные помещения - учебные аудитории для проведения занятий лекционного типа, лабораторных работ, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения на семестр учебного года выделяются в соответствии с расписанием занятий.

Специальные помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории (мультимедийным проектором, экраном, либо свободным участком стены ровного светлого тона размером не менее 2х1.5 метра, стандартной доской для работы с маркером). В случае отсутствия стационарной установки аудитория оснащена розетками электропитания для подключения переносного комплекта мультимедийной аппаратуры и экраном (либо свободным участком стены ровного светлого тона размером не менее 2х1.5 метра).

Для проведения занятий лекционного типа предлагаются демонстрационное оборудование и учебно-наглядные пособия, обеспечивающие тематические иллюстрации, соответствующие рабочей учебной программе дисциплины.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации.

|  |  |  |
| --- | --- | --- |
| Разработчик программы, профессор | \_\_\_\_\_\_\_\_\_\_ | Н.В.Грибкова |
| «\_19\_» \_января\_ 2018 г. |  |  |