

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Антипова Наталья Викторовна
Должность: и.о. директора филиала
Дата подписания: 20.02.2024 18:49:52
Уникальный программный ключ:
fae5412acb1bf810dc69e6bc004ac45622b84b3a

Приложение 3
к основной профессиональной образовательной программе
по направлению подготовки 38.03.01. «Экономика»
направленность (профиль) программы «Бизнес статистика и аналитика»

**Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования**

«Российский экономический университет имени Г.В. Плеханова»

Улан-Баторский филиал РЭУ им. Г.В. Плеханова



Одобрено
на заседании Совета Улан-Баторского
филиала РЭУ им Г.В. Плеханова
Протокол от № 9 от 02 июня 2022
Председатель Совета
Н.В. Антипова

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.18 Семантический анализ и текстовые данные

Направление подготовки: 38.03.01 Экономика

Направленность (профиль) программы: «Бизнес статистика и аналитика»

Уровень высшего образования Бакалавриат

Год начала подготовки 2022

Улан-Батор – 2022 г

Составитель(и):

К.э.н., доц., доцент
(ученая степень, ученое звание, должность,)

А.В. Безруков

Рабочая программа утверждена на заседании кафедры статистики
протокол № 8 от «29» апреля 2022 г.

СОДЕРЖАНИЕ

I. ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЙ РАЗДЕЛ	4
Цель и задачи освоения дисциплины	4
Место дисциплины в структуре образовательной программы	4
Объем дисциплины и виды учебной работы	4
Перечень планируемых результатов обучения по дисциплине	5
II. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ	6
III. УЧЕБНО-МЕТОДИЧЕСКОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ	8
РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА	8
ПЕРЕЧЕНЬ ИНФОРМАЦИОННО-СПРАВОЧНЫХ СИСТЕМ	9
ПЕРЕЧЕНЬ ЭЛЕКТРОННО-ОБРАЗОВАТЕЛЬНЫХ РЕСУРСОВ	9
ПЕРЕЧЕНЬ ПРОФЕССИОНАЛЬНЫХ БАЗ ДАННЫХ	9
ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ "ИНТЕРНЕТ", НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ	9
ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ	9
МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ	10
IV. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ	11
V. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ОПРЕДЕЛЯЮЩИЕ ПРОЦЕДУРЫ ОЦЕНИВАНИЯ ЗНАНИЙ И УМЕНИЙ, ХАРАКТЕРИЗУЮЩИХ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ	11
VI. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ	12

I. ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЙ РАЗДЕЛ

Цель и задачи освоения дисциплины

Целью дисциплины «Семантический анализ и текстовые данные» является приобретение студентами необходимой квалификации для реализации семантического анализа данных, в том числе с использованием программной среды Python

Задачами дисциплины являются:

1. Изучение концептуальной направленности методов машинного обучения как целевого применения статистических инструментов
2. Изучение методов направления «контролируемого» машинного обучения, поиска закономерностей по начальным условиям
3. Изучение методов направления «неконтролируемого» машинного обучения, поиска закономерностей без начальных условий
4. Изучение методов нейросетевого моделирования, укрепления, основ глубокого обучения.

Место дисциплины в структуре образовательной программы

Дисциплина «Семантический анализ и текстовые данные», относится к части учебного плана, формируемая участниками образовательных отношений.

Объем дисциплины и виды учебной работы

Таблица 1

Показатели объема дисциплины	Всего часов по формам обучения		
	очная	очно-заочная	заочная
Объем дисциплины в зачетных единицах	4 ЗЕТ		
Объем дисциплины в акад. часах	144		
Промежуточная аттестация: форма	Экз	Экзамен	-
Контактная работа обучающихся с преподавателем (Контакт. часы), всего:		30	-
1. Аудиторная работа (Ауд.), акад. часов всего, в том числе:	48	26	-
● лекции	18	10	-
● практические занятия	30	16	-
● лабораторные занятия	-	-	-
в том числе практическая подготовка	-	-	-
2. Индивидуальные консультации (ИК)**(заполняется при наличии по дисциплине курсовых работ/проектов)		-	-
3. Контактная работа по промежуточной аттестации (Катт) (заполняется при	2	2	-

<i>наличии по дисциплине курсовых работ/проектов)</i>			
4. Консультация перед экзаменом (КЭ)	2	2	-
5. Контактная работа по промежуточной аттестации в период экз. сессии / сессии заочников (Каттэк)		-	-
Самостоятельная работа (СР), всего:	92	114	-
в том числе:		-	
● самостоятельная работа в период экз. сессии (СРэк) (<i>заполняется при наличии экзамена по дисциплине</i>)	-	-	-
● самостоятельная работа в семестре (СРс)	-	114	-
в том числе, самостоятельная работа на курсовую работу (<i>заполняется при наличии по дисциплине курсовых работ/проектов</i>)	-	-	-
● изучение ЭОР (<i>при наличии</i>)	-	-	-
● изучение онлайн-курса или его части	-	-	-
● выполнение индивидуального или группового проекта	-	-	-

Перечень планируемых результатов обучения по дисциплине

Таблица 2

Формируемые компетенции (код и наименование компетенции)	Индикаторы достижения компетенций (код и наименование индикатора)	Результаты обучения (знания, умения)
<i>ПК-2. Способен проводить процедуру сводки статистических данных по утвержденным методикам</i>	<i>ПК-2.1. Осуществляет формирование входных массивов информации баз данных</i>	<i>ПК-2.1. 3-1. Знает методические документы по формированию входных массивов статистических данных</i>
		<i>ПК-2.1. 3-2. Знает нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных</i>
		<i>ПК-2.1. У-1. Умеет формировать входные массивы статистических данных</i>
		<i>ПК-2.1. У-2. Умеет контролировать сохранность статистической информации</i>
	<i>ПК-2.3. Формирует выходные массивы информации</i>	<i>ПК-2.3. 3-1. Знает инструкции по формированию выходных массивов статистических данных</i>
		<i>ПК-2.3. 3-2. Знает инструкции по осуществлению логического и арифметического контроля</i>
		<i>ПК-2.3. 3-3. Знает нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных</i>

		<i>ПК-2.3. У-1. Умеет формировать выходные массивы статистической информации</i>
		<i>ПК-2.3. У-2. Умеет осуществлять логический и арифметический контроль выходной информации</i>
		<i>ПК-2.3. У-3. Умеет контролировать сохранность статистической информации</i>

II. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ
этапы формирования и критерии оценивания сформированности компетенций

Таблица 3

№ п/п	Наименование раздела, темы дисциплины	Трудоемкость, академические часы						Индикаторы достижения компетенций	Результаты обучения (знания, умения)	Учебные задания для аудиторных занятий	Текущий контроль	Задания для творческого рейтинга (по теме(-ам)/разделу или по всему курсу в целом)
		Лекции	Практические занятия	Лабораторные занятия	Практическая подготовка	Самостоятельная работа	Всего					
Семестр 3												
1.	Введение в дисциплину. Конвейер интеллектуального анализа текста: поиск информации, извлечение информации и интеллектуальный анализ данных	1	2	0	0	2	5	ПК-2.1. ПК-2.3	ПК-2.1. 3-1 ПК-2.1. 3-2 ПК-2.1. У-1 ПК-2.1. У-2 ПК-2.3. 3-1 ПК-2.3. 3-2 ПК-2.3. 3-3 ПК-2.3. У-1 ПК-2.3. У-2 ПК-2.3. У-3	О.	Т.	-
2.	Статистическая обработка естественного языка Основы обработки естественного языка: лингвистические основы, уровни лингвистического анализа.	1	2	0	0	4	7	ПК-2.1. ПК-2.3	ПК-2.1. 3-1 ПК-2.1. 3-2	О.	Т.	-

									ПК-2.1. У-1 ПК-2.1. У-2 ПК-2.3. 3-1 ПК-2.3. 3-2 ПК-2.3. 3-3 ПК-2.3. У-1 ПК-2.3. У-2 ПК-2.3. У-3			
3.	Анализ соответствий. Латентное семантическое индексирование Подходы к интеллектуальному анализу текста: основанные на правилах против машинного обучения против гибридных; общий или предметно-ориентированный; адаптация домена. Работа с реальным текстом: типы текста, форматы документов и преобразование, кодировки символов, разметка, низкоуровневые процессы (разделение предложений, токенизация, тегирование части речи, разбиение на фрагменты).						ПК-2.1. ПК-2.3	ПК-2.1. 3-1 ПК-2.1. 3-2 ПК-2.1. У-1 ПК-2.1. У-2 ПК-2.3. 3-1 ПК-2.3. 3-2 ПК-2.3. 3-3 ПК-2.3. У-1 ПК-2.3. У-2 ПК-2.3. У-3				
4.	«Связывание сущностей» извлечение терминов, распознавание именованных сущностей, извлечение отношений, извлечение фактов и событий; частичный анализ против полного анализа.						ПК-2.1. ПК-2.3	ПК-2.1. 3-1 ПК-2.1. 3-2 ПК-2.1. У-1 ПК-2.1.				

									У-2 ПК-2.3. 3-1 ПК-2.3. 3-2 ПК-2.3. 3-3 ПК-2.3. У-1 ПК-2.3. У-2 ПК-2.3. У-3			
5.	Анализ текстовых данных Интеллектуальный анализ данных и визуализация результатов интеллектуального анализа текста. Оценка систем интеллектуального анализа текста: меры оценки, роль задач оценки, оценка юзабилити.							ПК-2.1. ПК-2.3	ПК-2.1. 3-1 ПК-2.1. 3-2 ПК-2.1. У-1 ПК-2.1. У-2 ПК-2.3. 3-1 ПК-2.3. 3-2 ПК-2.3. 3-3 ПК-2.3. У-1 ПК-2.3. У-2 ПК-2.3. У-3			
6.	Анализ социальных сетей Ресурсы для интеллектуального анализа текста: аннотированные корпуса, вычислительная лексика, онтологии, вычислительные грамматики; вопросы проектирования, строительства и эксплуатации. Проблемы крупномасштабной обработки текста: распределенный интеллектуальный анализ текста, масштабируемые системы интеллектуального анализа текста.							ПК-2.1. ПК-2.3	ПК-2.1. 3-1 ПК-2.1. 3-2 ПК-2.1. У-1 ПК-2.1. У-2 ПК-2.3.			

									3-1 ПК-2.3. 3-2 ПК-2.3. 3-3 ПК-2.3. У-1 ПК-2.3. У-2 ПК-2.3. У-3			
7.	Основы управления знанием Сэмплер приложений и сервисов интеллектуального анализа текста; тематические исследования.							ПК- 2.1. ПК-2.3	ПК-2.1. 3-1 ПК-2.1. 3-2 ПК-2.1. У-1 ПК-2.1. У-2 ПК-2.3. 3-1 ПК-2.3. 3-2 ПК-2.3. 3-3 ПК-2.3. У-1 ПК-2.3. У-2 ПК-2.3. У-3			
	Итого	18	30	0	0	92	140					

III. УЧЕБНО-МЕТОДИЧЕСКОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Основная литература (О):

1. Анализ данных: учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва: Издательство Юрайт, 2020. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст: электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450166>
2. Миркин, Б. Г. Введение в анализ данных: учебник и практикум / Б. Г. Миркин. — Москва: Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст: электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450262>

Дополнительная литература (Д):

1. Черткова, Е. А. Статистика. Автоматизация обработки информации: учебное пособие для вузов / Е. А. Черткова; под общей редакцией Е. А. Чертковой. — 2-е изд., испр. и доп. — Москва: Издательство Юрайт, 2020. — 195 с. — (Высшее образование). — ISBN 978-5-534-01429-7. — Текст: электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/452447>
2. Иванов, Ю. Н. Экономическая статистика: учебник / под ред. Ю.Н. Иванова. — 5-е изд., перераб. и доп. — Москва: ИНФРА-М, 2020. — 584 с. + Доп. материалы [Электронный ресурс]. — (высшее образование: Бакалавриат). — DOI 10.12737/7728. - ISBN 978-5-16-106497-9. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1063310>
3. Дубина, И. Н. Основы математического моделирования социально-экономических процессов: учебник и практикум для вузов / И. Н. Дубина. — Москва: Издательство Юрайт, 2020. — 349 с. — (Высшее образование). — ISBN 978-5-534-00501-1. — Текст: электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450960>
4. Королев, А. В. Экономико-математические методы и моделирование: учебник и практикум для вузов / А. В. Королев. — Москва: Издательство Юрайт, 2020. — 280 с. — (Высшее образование). — ISBN 978-5-534-00883-8. — Текст: электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/451297>

ПЕРЕЧЕНЬ ИНФОРМАЦИОННО-СПРАВОЧНЫХ СИСТЕМ

1. <http://www.consultant.ru> - Справочно-правовая система Консультант Плюс;
2. <http://www.garant.ru> - Справочно-правовая система Гарант.

ПЕРЕЧЕНЬ ЭЛЕКТРОННО-ОБРАЗОВАТЕЛЬНЫХ РЕСУРСОВ

1. Садовникова Н.А. Теория статистики. *(электронный образовательный ресурс, размещён в ЭОС РЭУ им. Г.В. Плеханова)* <http://lms.rea.ru/>

ПЕРЕЧЕНЬ ПРОФЕССИОНАЛЬНЫХ БАЗ ДАННЫХ

1. Федеральная служба государственной статистики (Росстат) — <http://www.gks.ru/>
2. Единая межведомственная информационно-статистическая система (ЕМИСС) —
3. <https://fedstat.ru/>
4. Всемирный банк — <https://data.worldbank.org/>

ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ "ИНТЕРНЕТ", НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1. Ситуационный центр Российского экономического университета имени Г. В. Плеханова — <http://reu.stat.university:8180/biportal/contourbi.jsp/>

ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

№ п/п	Перечень информационных технологий, программного обеспечения
1	Отечественная операционная система
2	Прикладной пакет документов

МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Дисциплина «Семантический анализ и текстовые данные» обеспечена:

для проведения занятий лекционного типа:

1. учебной аудиторией, оборудованной учебной мебелью, мультимедийными средствами обучения для демонстрации лекций-презентаций;

для проведения занятий семинарского типа (практические занятия, лабораторные занятия):

1. учебной аудиторией, оборудованной учебной мебелью и техническими средствами обучения, служащими для представления учебной информации: проектор, ПК, интерактивная-доска; для самостоятельной работы, в том числе для курсового проектирования:

1. помещением для самостоятельной работы, оснащенным компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде университета.

IV. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

- Положение о курсовых работах (проектах) в ФГБОУ ВО "РЭУ им. Г.В. Плеханова".
- Методические рекомендации по составлению и оформлению междисциплинарного проекта.
- Методические рекомендации по организации и выполнению внеаудиторной самостоятельной работы.
- Методические указания по подготовке и оформлению рефератов.

V. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ОПРЕДЕЛЯЮЩИЕ ПРОЦЕДУРЫ ОЦЕНИВАНИЯ ЗНАНИЙ И УМЕНИЙ, ХАРАКТЕРИЗУЮЩИХ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ

Результаты текущего контроля и промежуточной аттестации формируют рейтинговую оценку работы обучающегося. Распределение баллов при формировании рейтинговой оценки работы обучающегося осуществляется в соответствии с «Положением о рейтинговой системе оценки успеваемости и качества знаний студентов в процессе освоения дисциплины «Статистическая семантика в машинном обучении» в федеральном государственном бюджетном образовательном учреждении высшего образования «Российский экономический университет имени Г.В. Плеханова».

Таблица 4

Виды работ	Максимальное количество баллов
Выполнение учебных заданий на аудиторных занятиях	20
Текущий контроль	20
Творческий рейтинг	20
Промежуточная аттестация (экзамен)	40
ИТОГО	100

В соответствии с Положением о рейтинговой системе оценки успеваемости и качества знаний обучающихся «преподаватель кафедры, непосредственно ведущий занятия со студенческой группой, обязан проинформировать группу о распределении рейтинговых баллов по всем видам работ на первом занятии учебного модуля (семестра), количестве модулей по учебной дисциплине, сроках и формах

контроля их освоения, форме промежуточной аттестации, снижении баллов за несвоевременное выполнение выданных заданий. Обучающиеся в течение учебного модуля (семестра) получают информацию о текущем количестве набранных по дисциплине баллов через личный кабинет студента».

VI. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ¹

Оценочные средства по дисциплине разработаны в соответствии с Положением о фонде оценочных средств в федеральном государственном бюджетном образовательном учреждении высшего образования «Российский экономический университет имени Г.В. Плеханова».

Типовой перечень вопросов к экзамену:

1. Модели и их свойства. Аналитический и информационный подходы к моделированию.
 2. Формы представления, типы и виды анализируемых данных.
 3. Обучение моделей «с учителем» и «без учителя». Обучающее и тестовое множество. Ошибки обучения. Эффект переобучения.
 4. Общая схема анализа данных. Требования к алгоритмам анализа данных.
 5. Характеристика этапов технологии KDD.
 6. Data Mining. Характеристика классов задач, решаемых методами Data Mining.
 7. Программный инструментарий для выполнения анализа данных.
 8. Основные положения концепции хранилищ данных (DW).
-
1. Цели и задачи аффинитивного анализа. Поддержка и достоверность ассоциативных правил. Лифт и леввередж.
 2. Сферы применения ассоциативных правил.
 3. Иерархические ассоциативные правила.
 4. Цели, задачи и основное содержание кластерного анализа. Классификация методов кластеризации.
 5. Способы определения меры расстояния между кластерами.
 6. Характеристика методов связи для процедуры кластеризации (одиночная, полная, средняя).
 7. Алгоритм кластеризации k-means.
 8. Сети Кохонена (KCN).
 9. Карты Кохонена (SOM).

¹ В данном разделе приводятся примеры оценочных средств

10. Проблемы алгоритмов кластеризации.
11. Цели, задачи и отличительные особенности классификации и регрессии.
12. Сферы применения методов классификации и регрессии.
13. Простая линейная регрессия.
14. Оценка соответствия простой линейной регрессии реальным данным.
15. Простая регрессионная модель.
16. Оценка значимости простой регрессионной модели (t-критерий и F-критерий).
17. Множественная линейная регрессия.
18. Модель множественной линейной регрессии.
19. Оценка значимости множественной регрессионной модели.
20. Регрессия с категориальными входными переменными.
21. Методы отбора переменных в регрессионные модели.
22. Ограничения применимости регрессионных моделей.
23. Логистическая регрессия. Интерпретация модели логистической регрессии.
24. Множественная логистическая регрессия.
25. Цели, задачи и принципы построения деревьев решений. Общая характеристика алгоритмов построения деревьев решений.
26. Сферы применения деревьев решений.
27. Алгоритмы ID3 и C4.5.
28. Алгоритм CART.
29. Упрощение деревьев решений.
30. Цели, задачи и принципы работы нейронных сетей.
31. Принципы функционирования многослойного персептрона.
32. Алгоритмы обучения нейронных сетей.
33. Алгоритм обратного распространения ошибки.
34. Общая характеристика временных рядов и их компонентов. Цели и задачи анализа временных рядов.
35. Цели, задачи и принципы прогнозирования. Модели прогнозирования. Обобщенная модель прогноза.

Типовые тестовые задания:

- 1) Какой метод используется для оценки параметров уравнения используется, в соответствии с которым $\hat{\beta} = (X'X)^{-1}X'Y$, где $\hat{\beta}$ – вектор оценок параметров уравнения, X – матрица независимых переменных, Y – вектор независимых переменных.
 1. Метод k-средних;
 2. Метод наименьших квадратов;
 3. Индексный метод;
 4. Иерархические кластер-процедуры.
- 2) По какой формуле определяется наблюдаемое значение F-критерия Фишера (R^2 – множественный коэффициент детерминации, n – число наблюдений, m – число независимых переменных)?

1. $F_{\text{набл}} = \frac{R}{1-R^2} \times \frac{n-m}{m}$
2. $F_{\text{набл}} = \frac{R^2}{1-R^2} \times \frac{n}{m}$
3. $F_{\text{набл}} = \frac{R^2}{1-R^2} \times \frac{n-m-1}{m}$
4. $F_{\text{набл}} = \frac{R^2}{1-R^2} \times \frac{n-1}{m}$

3) Для оценки надежности регрессионного уравнения наблюдаемое значение F -критерия Фишера должно быть:

1. Больше расчетного значения F -критерия Фишера
2. Меньше расчетного значения F -критерия Фишера
3. Равно расчетному значению F -критерия Фишера
4. F -критерий Фишера не оценивает надежность уравнения.

4) Для определения количества значений, находящихся в выборке ниже или выше медианного значения, используется:

1. Коэффициент конкордации;
2. Коэффициент корреляции рангов Спирмена;
3. Коэффициент корреляции рангов Кэндалла;
4. Критерий знаков.

5) Какая математическая модель приведена ниже?

$x_{ij} = a_{i1}f_{1j} + a_{i2}f_{2j} + \dots + a_{im}f_{mj} + d_jv_{ij}$, где $i = \underline{1, n}$; $j = \underline{1, m}$; $r = \underline{1, m}$, где n - число объектов наблюдения; m - число показателей, характеризующих объект; r - число значимых общих факторов; x_{ij} - центрированное значение j -го показателя (переменной) у i -го объекта исследования; f_r - r -й общий фактор; v_j - j -й характерный (индивидуальный) фактор, присущий только данной j -й переменной; a_{jr} - весовой коэффициент j -й переменной на r -м общем факторе; d_j - весовой коэффициент j -й переменной на j -м характерном факторе.

1. Модель факторного анализа;
2. Математическая модель метода главных компонент;
3. Мультипликативная модель;
4. Аддитивная модель.

Типовые расчетно-аналитические задания/задачи:

Тема 2: Статистическая обработка естественного языка

1. По данным провести классификацию предприятий и построить дендрограмму.
2. По дендрограмме выбрать окончательный вариант разбиения предприятий на кластеры и дать интерпретацию полученным результатам, используя статистические характеристики кластеров.

№ варианта	результативный признак, у	номера факторных признаков, х
1	1	6,8,11,12,17
2	2	6,8,11,13,17
3	3	8,11,12,13,17
4	1	6,8,13,14,17
5	2	8,11,13,14,17

Таблица 2.

№ п/п	Y ₁	Y ₂	Y ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀
1	9,26	204,2	13,26	0,23	0,78	0,40	1,37	1,23	0,23	1,45
2	9,38	209,6	10,16	0,24	0,75	0,26	1,49	1,04	0,39	1,30
3	12,11	222,6	13,72	0,19	0,68	0,40	1,44	1,80	0,43	1,37
4	10,81	236,7	12,85	0,17	0,70	0,50	1,42	0,43	0,18	1,65
5	9,35	62,0	10,63	0,23	0,62	0,40	1,35	0,88	0,15	1,91

№ п/п	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	X ₁₇
1	26006	167,69	47750	6,40	166,32	10,08	17,72
2	23935	186,1	50391	7,80	92,88	14,76	18,39
3	22589	220,45	43149	9,76	158,04	6,48	26,46
4	21220	169,3	41089	7,90	93,96	21,96	22,37
5	7394	39,53	14257	5,35	173,88	11,88	28,13

Y₁- производительность труда;

Y₂- индекс снижения себестоимости продукции;

Y₃ -рентабельность;

X₄-трудоемкость единицы продукции;

X₅ -удельный вес рабочих;

X₆- удельный вес покупных изделий;

X₇-коэффициент сменности оборудования;

X₈- премии и вознаграждения на одного работника;

X₉- удельный вес потерь от брака;

X₁₀- фондоотдача;

X₁₁- среднегодовая численность промышленно-производственного персонала;

X₁₂- среднегодовая стоимость ОПФ;

X₁₄-фондовооруженность;

X₁₅- оборачиваемость нормируемых оборотных средств;

X₁₆- оборачиваемость ненормируемых оборотных средств;

X₁₇-непроизводственные расходы.

Типовая структура экзаменационного билета/зачетного задания

<i>Наименование оценочного средства</i>	<i>Максимальное количество баллов</i>
1. <i>Вопрос 1.</i> Общая теория многомерных распределений.	5
<i>Вопрос 2.</i> Оценки частных и множественных коэффициентов корреляции	8
<i>Практическое задание (расчетно-аналитическое)</i> В таблице представлены данные о количестве заключенных договоров в разрезе страховых продуктов каждым агентом. Используя статистический ППП, по представленным данным необходимо: <ol style="list-style-type: none"> 1) Построить интервальный вариационный ряд, гистограмму. 2) Вычислить выборочные характеристики: среднее значение, дисперсию, СКО, коэффициент вариации, асимметрию, эксцесс, моду, медиану. 3) Проверить гипотезу о нормальном законе распределения анализируемой случайной величины. 4) Построить 95% интервальные оценки математического ожидания, дисперсии и СКО. 	12

Страховой агент, №	Страхование имущества	Ипотечное страхование	Автострахование	Страхование туристов	ОСАГО	ДМС	Всего
1	26	10	6	4	11	7	64
2	38	18	17	3	10	2	86
3	62	4	5	5	9	1	87
4	56	24	26	3	4	8	122
5	19	20	21	1	6	6	73
6	85	8	7	3	21	7	130
7	44	9	2	1	14	8	78
8	63	10	9	0	5	5	93
9	52	7	7	4	8	6	84
10	47	5	1	1	9	10	73
11	57	10	9	5	10	1	92
12	28	13	12	1	12	2	68
13	20	4	5	1	8	7	45
14	70	18	18	3	16	3	128
15	46	17	17	0	10	5	95
16	30	13	13	1	13	3	73
17	38	15	14	2	9	14	93
18	38	17	17	1	10	12	96
19	53	15	15	1	15	10	109
20	48	15	14	1	9	9	95
21	57	14	13	2	6	21	113
22	62	26	29	1	13	2	134
23	46	10	9	1	3	6	74

24	59	27	30	4	10	11	142
25	30	18	17	3	21	15	103
26	68	10	9	2	7	4	100
27	61	13	0	1	11	7	93
28	56	5	6	1	1	12	82
29	39	4	1	2	10	7	63
30	53	11	10	3	7	7	91
31	59	16	15	1	11	6	107
32	39	0	3	2	6	14	64
33	40	14	13	1	9	5	82
34	59	11	10	1	12	3	96
35	47	0	3	4	6	9	70
36	22	31	38	5	13	7	115
37	63	17	16	4	12	7	119
38	53	1	3	2	18	10	88
39	48	12	10	1	1	9	82
40	57	23	24	0	11	12	127
41	39	17	16	4	12	4	91
42	64	8	0	4	10	2	88
43	47	10	9	5	7	5	84
44	40	5	2	4	12	6	69
45	23	12	11	1	6	13	66
46	27	3	4	0	2	8	44
47	43	12	11	4	12	5	88
48	72	19	19	2	19	1	132
49	40	21	22	2	13	4	104
50	39	16	16	2	16	14	103

Практическое задание (расчетно-аналитическое)

Для оценки удовлетворенности населения жизнью проведен социологический опрос в 80 странах мира. В таблице представлены агрегированные результаты проведенного опроса в разрезе стран, единицей измерения показателей является удельный вес респондентов, положительно оценивающих параметр. В качестве основных параметров удовлетворенности жизнью выбраны следующие:

x_1 – работа, %;

x_2 – здоровье, %;

x_3 - материальное благополучие, %;

x_4 - достижение поставленных целей, %;

x_5 - социальный статус, %;

x_6 - социальные контакты, %.

Необходимо с помощью метода главных компонент с помощью ППП STATISTICA выявить факторы удовлетворенности населения жизнью и ранжировать страны по уровню удовлетворенности населения жизнью.

Результаты опроса населения стран об удовлетворенности жизнью

№	Страна	Показатели					
		x_1	x_2	x_3	x_4	x_5	x_6
1	Израиль	80	80	71	88	81	85

15

2	Греция	80	82	57	90	92	79
3	Словакия	76	72	47	85	78	93
4	Эстония	79	64	46	72	79	85
5	Венгрия	83	69	43	88	88	90
6	Португалия	90	80	47	92	93	87
7	Польша	82	72	67	87	91	89
8	Чили	81	73	68	90	93	83
9	Латвия	79	63	33	79	80	78
10	Хорватия	78	77	48	83	74	90
11	Болгария	73	67	29	77	77	81
12	Тринидад и Тобаго	76	82	40	97	93	85
13	Сербия	73	73	35	84	77	82
14	Белоруссия	66	55	34	70	71	88
15	Перу	74	72	54	96	89	79
16	Албания	72	75	43	78	68	79
17	Россия	74	56	36	79	83	88
18	Казахстан	82	68	51	88	81	88
19	Азербайджан	73	68	42	87	79	72
20	Босния и Герцеговина	76	75	39	80	67	74
21	Украина	71	55	23	74	78	81
22	Иран	71	82	55	87	81	62
23	Югославия	71	82	34	93	81	78
24	Грузия	63	50	22	86	83	54
25	Армения	61	53	31	93	89	67
26	Эквадор	80	76	57	98	93	78
27	Гайана	79	83	69	90	75	83
28	Ямайка	82	88	50	98	80	91
29	Турция	71	76	44	85	68	84
30	Доминиканская Республика	69	80	57	96	92	84
31	Сальвадор	82	80	60	97	89	72
32	Шри-Ланка	86	77	58	91	76	82
33	Таиланд	91	79	63	95	75	82
34	Боливия	83	79	67	94	90	82
35	Парагвай	85	84	63	93	96	89
36	Филиппины	83	77	68	96	94	77
37	Ботсвана	58	67	41	92	83	83
38	Молдова	68	60	39	79	73	83
39	Гана	84	87	64	95	77	84
40	Намбия	84	87	61	98	86	83
41	Гондурас	84	83	65	95	91	81
42	Индонезия	63	83	62	95	92	78
43	Киргизия	78	74	48	91	86	85
44	ЮАР	66	79	42	97	83	88
45	Таджикистан	78	75	69	91	76	65
46	Вьетнам	72	79	59	98	92	79
47	Марокко	69	88	71	90	89	85
48	Никарагуа	80	80	62	98	91	83
49	Индия	74	85	61	91	72	66
50	Камбоджи	80	69	51	81	87	82
51	Кения	57	70	25	98	78	79
52	Бангладеш	76	73	63	94	87	53
53	Гана	54	66	34	98	88	63
54	Камерун	63	69	40	93	85	73
55	Йемен	74	80	53	88	84	75
56	Мадагаскар	46	76	24	96	77	77
57	Мавритания	57	79	47	93	85	81
58	Нигерия	65	80	40	92	81	72
59	Уганда	53	64	35	96	79	85
60	Сенегал	39	68	27	89	85	81
61	Гаити	51	51	35	81	66	64
62	Ангола	72	67	54	90	83	58
63	Танзания	45	67	21	95	74	76
64	Замбия	48	78	34	93	83	62
65	Руанда	41	64	37	88	77	56
66	Малави	62	77	64	99	88	72
67	Судан	65	77	64	97	89	89
68	Афганистан	71	79	53	83	64	54
69	Гвинея	68	75	27	96	86	58

70	Эфиопия	50	79	33	89	74	76
71	Сьерра Леон	49	47	19	98	81	59
72	Центрально-Африканская республика	78	81	31	96	74	56
73	Мали	30	71	30	99	86	75
74	Буркина-Фасо	46	70	27	94	83	73
75	Либерия	47	70	46	100	82	58
76	Чад	78	69	52	93	79	57
77	Мозамбик	74	82	46	93	89	75
78	Нигер	54	82	52	99	93	77
79	Конго	60	74	40	98	79	67
80	Зимбабве	49	72	27	91	81	81

Показатели и критерии оценивания планируемых результатов освоения компетенций и результатов обучения, шкала оценивания

Таблица 5

Шкала оценивания		Формируемые компетенции	Индикатор достижения компетенции	Критерии оценивания	Уровень освоения компетенций
85 – 100 баллов	«отлично»/ «зачтено»	ПК-2. Способен проводить процедуру сводки статистических данных по утвержденным методикам	ПК-2.1. Осуществляет формирование входных массивов информации баз данных	Знает верно и в полном объеме: методические документы по формированию входных массивов статистических данных, нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных Умеет верно и в полном объеме: формировать входные массивы статистических данных, контролировать сохранность статистической информации	Продвинутый
			ПК-2.3. Формирует выходные массивы информации	Знает верно и в полном объеме: инструкции по формированию выходных массивов статистических данных, инструкции по осуществлению логического и арифметического контроля, нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных Умеет верно и в полном объеме: формировать выходные массивы статистической информации, осуществлять логический и арифметический контроль выходной информации, контролировать сохранность статистической информации	

70 – 84 баллов	«хорошо»/ «зачтено»	<i>ПК-2. Способен проводить процедуру сводки статистических данных по утвержденным методикам</i>	<i>ПК-2.1. Осуществляет формирование входных массивов информации баз данных</i>	Знает с незначительными замечаниями: методические документы по формированию входных массивов статистических данных, нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных Умеет с незначительными замечаниями: формировать входные массивы статистических данных, контролировать сохранность статистической информации	Повышенный
			<i>ПК-2.3. Формирует выходные массивы информации</i>	Знает с незначительными замечаниями: инструкции по формированию выходных массивов статистических данных, инструкции по осуществлению логического и арифметического контроля, нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных Умеет с незначительными замечаниями: формировать выходные массивы статистической информации, осуществлять логический и арифметический контроль выходной информации, контролировать сохранность статистической информации	
50 – 69 баллов	«удовлетворительно»/ «зачтено»	<i>ПК-2. Способен проводить процедуру сводки статистических данных по утвержденным методикам</i>	<i>ПК-2.1. Осуществляет формирование входных массивов информации баз данных</i>	Знает на базовом уровне, с ошибками: методические документы по формированию входных массивов статистических данных, нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных Умеет на базовом уровне, с ошибками: формировать входные массивы статистических данных, контролировать сохранность статистической информации	Базовый
			<i>ПК-2.3. Формирует выходные массивы информации</i>	Знает на базовом уровне, с ошибками: инструкции по формированию выходных массивов статистических данных, инструкции по осуществлению логического и арифметического контроля, нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных Умеет на базовом уровне, с ошибками: формировать выходные массивы статистической информации, осуществлять логический и арифметический контроль выходной информации, контролировать сохранность статистической информации	

менее 50 баллов	«неудовлетворительно»/ «не зачтено»	ПК-2. Способен проводить процедуру сводки статистических данных по утвержденным методикам	ПК-2.1. Осуществляет формирование входных массивов информации баз данных	<p>Не знает на базовом уровне: методические документы по формированию входных массивов статистических данных, нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных</p> <p>Не умеет на базовом уровне: формировать входные массивы статистических данных, контролировать сохранность статистической информации</p>	Компетенции не сформированы
			ПК-2.3. Формирует выходные массивы информации	<p>Не знает на базовом уровне: инструкции по формированию выходных массивов статистических данных, инструкции по осуществлению логического и арифметического контроля, нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных</p> <p>Не умеет на базовом уровне: формировать выходные массивы статистической информации, осуществлять логический и арифметический контроль выходной информации, контролировать сохранность статистической информации</p>	

