

*Федеральное государственное автономное образовательное учреждение высшего образования «Российский университет дружбы народов»*

*Высшая школа промышленной политики и предпринимательства*

Рекомендовано МССН

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

«Управление процессами на основе больших данных»

**Направление подготовки: 38.04.01.ЭКОНОМИКА**

**Направленность программы (специализации) «Принятие эффективных управленческих решений (Big Data Economics)»**

**Москва – 2021**

## 1. Цель и задачи дисциплины.

Цель учебной дисциплины заключается в получении обучающимися теоретических знаний о процессах управления с последующим применением в профессиональной сфере и развитии практических навыков в области науки об управлении, с последующим применением в профессиональной сфере, формирование способностей решать средствами математики задачи управления.

Углубить изучение методологических и теоретических проблем, связанных с выявлением устойчивых, повторяющихся связей в социально-экономических процессах, их структурных характеристик, закономерностей функционирования и тенденций развития экономических отношений, объяснением на этой основе существующих факторов и феноменов социально-экономической жизни, пониманием и предвидением хозяйственно-политических событий.

Задачи учебной дисциплины:

1. Усвоение знаний о сущности, структуре и видах математических моделей принятия решений;
2. Формирование представлений о содержании, формах, особенностях дисциплины «Управление процессами на основе больших данных»;
3. Развитие навыков создания и решения моделей, необходимых в сфере управления.

## 2. Место учебной дисциплины в структуре ОП ВО.

Учебная дисциплина «Управление процессами на основе больших данных» реализуется в вариативной части основной профессиональной образовательной программы высшего образования - программы подготовки научно-педагогических кадров в аспирантуре по направлению подготовки 09.06.01 «Информатика и вычислительная техника», профилю подготовки «Управление в социальных и экономических системах» очной и заочной формам обучения.

Изучение учебной дисциплины «Технология работы с большими данными» является базовым для последующего освоения программного материала учебных дисциплин: «Методы научных исследований».

Курс построен таким образом, чтобы обеспечить вклад в фундаментальную подготовку студентов для осуществления ими научно-исследовательской и аналитической деятельности, а также педагогической деятельности в высших и средних профессиональных учебных заведениях.

Процесс освоения учебной дисциплины направлен на формирование у обучающихся следующих универсальных и общепрофессиональных компетенций: ОПК 2, УК 7, в соответствии с основной профессиональной образовательной программой высшего образования - программой подготовки научно-педагогических кадров в аспирантуре по направлению подготовки 09.06.01 «Информатика и вычислительная техника», профилю подготовки «Управление в социальных и экономических системах».

## 3. Требования к результатам освоения дисциплины:

Процесс изучения дисциплины направлен на формирование следующих компетенций:

### *а) универсальных (УК)*

- способность осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода (УК-1);
- способность управлять проектом на всех этапах его жизненного цикла (УК-2);

**в) профессиональных (ПК)**

- способность представлять информацию бизнес-анализа различными способами и в различных форматах для обсуждения с заинтересованными сторонами (ПК-17);

В результате изучения дисциплины студент должен:

Код компетенции	Содержание компетенции	Результаты обучения
УК 7	способность оценивать влияние технологий больших данных на результаты решений исследовательских и практических задач, в том числе в междисциплинарных областях	Знать: математические методы обработки данных
		Уметь: выполнять исследования процессов создания, накопления и обработки информации, включая анализ и создание моделей данных и знаний, языков их описания и манипулирования
		Владеть: новыми методами исследования и обработки данных и их применению в самостоятельной научно-исследовательской деятельности в области профессиональной деятельности
ПКО 15	владение культурой научного исследования, в том числе с использованием современных информационно-коммуникационных технологий	Знать: методы и средства системного анализа, оптимизации, управления, принятия решений и обработки информации
		Уметь: разрабатывать методы проектирования и анализа алгоритмов, программ
		Владеть: современными методами исследования и информационно-коммуникационными технологиями

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

Общая трудоемкость учебной дисциплины для очной формы обучения составляет 3 зачетные единицы.

Вид учебной работы	Всего зачетных единиц	Курс			
		1	3/2	3	4
<b>Аудиторные учебные занятия, всего</b>	18		18		
В том числе:					
Учебные занятия лекционного типа	9		9		
Учебные занятия семинарского типа	9		9		
Лабораторные занятия					
<b>Самостоятельная работа обучающихся, всего</b>	45		45		
В том числе:					
Реферат			0		
<i>Другие виды самостоятельной работы:</i>					
Расчетное практическое задание					

<b>Вид промежуточной аттестации</b>	<b>Зачет</b>		<b>Зачет</b>		
<b>Контроль</b>	<b>9</b>		<b>9</b>		
<b>Общая трудоемкость учебной дисциплины</b>	<b>2</b>		<b>2</b>		
	<b>72</b>		<b>72</b>		

## 5. Содержание дисциплины

### 5.1. Содержание разделов дисциплины

Объем аудиторных занятий составляет

18час.

Объем самостоятельной работы – 45 часов.

№ п/ п	Раздел, тема	Виды учебной работы, академических часов						Формы текущего контроля успеваемости и промежуточной аттестации						
		Всего	Самостоят. работа	Аудиторные занятия				Конт. раб.	Реферат	доклад	Расчетное практическое задание	тестирование	Зачет	Экзамен (кандид.)
				Всего	Лекционного типа	Семинарского типа	Лабораторные занятия							
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1.	Методы многомерного статистического анализа и анализа нечисловой	54	29	26	9	18							*	
2.	Программные средства для обработки данных и системы Data Mining	54	28	25	8	16			*				*	

### 5.2. Темы учебной дисциплины и вопросы для самоподготовки

#### Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации

**Цель:** заключается в получении обучающимися теоретических знаний о принципах математических методов обработки данных, в соответствии с УК-7, с последующим применением в профессиональной сфере и формирование практических навыков по выполнению и исследованию процессов создания, накопления и обработки информации, включая анализ и создание моделей данных и знаний (УК-7), способности на этой основе самостоятельно осуществлять научно-исследовательскую деятельность (ОПК-2).

#### Перечень изучаемых элементов содержания

Простейшие статистические характеристики. Приведение к нормальной форме.

Оцифровка нечисловых данных. Предмет и содержание раздела «Многомерные статистические методы». Роль и сущность многомерных статистических методов в экономике, управлении, финансах, социальных науках: постановка основных задач, примеры практического использования в социально-экономических исследованиях.

Многомерное нормальное распределение как основная модель современных многомерных статистических методов. Практическое применение многомерных

методов в финансовых, экономических и социальных исследованиях. Методы статистического

оценивания многомерных параметров и проверки гипотез. Особенности анализа количественных и качественных признаков. Методы шкалирования. Кластерный анализ. Компонентный анализ. Факторный анализ.

#### **Вопросы для самоподготовки**

1. Показатели и объекты (измерения). Интервальные данные.
2. Нечисловые данные.
3. Простейшие статистические характеристики.
4. Приведение к нормальной форме.
5. Оцифровка нечисловых данных.
6. Роль и сущность многомерных статистических методов в экономике, управлении, финансах, социальных науках: постановка основных задач, примеры практического использования в социально-экономических исследованиях.
7. Многомерное нормальное распределение как основная модель современных многомерных статистических методов.
8. Практическое применение многомерных методов в финансовых, экономических и социальных исследованиях.
9. Методы статистического оценивания многомерных параметров и проверки гипотез.
10. Особенности анализа количественных и качественных признаков. Методы шкалирования.
11. Постановка основных прикладных задач классификации многомерных наблюдений.
12. Классификация с обучением и без обучения. Сущность методов классификации.
13. Кластерный анализ. Меры однородности объектов. Расстояния между объектами. Расстояния между кластерами.
14. Реализация методов кластерного анализа в современных пакетах прикладных программ.
15. Кластерный анализ финансовой деятельности предприятий.
16. Кластерный анализ мировой демографической статистики.
17. Кластерный анализ социологических опросов.
18. Кластерный анализ результатов аттестации персонала компании.
19. Зависимость выбора метода классификации от цели исследования.
20. Компонентный анализ. Математическая модель главных компонент. Геометрическая интерпретация главных компонент. Формирование названий главных компонент.
21. Экономическая интерпретация главных компонент.
22. Реализация методов компонентного анализа в современных пакетах прикладных программ.
23. Использование компонентного анализа в экономических и социальных исследованиях.
24. Линейная модель факторного анализа. Различие предпосылок компонентного и факторного анализа. Экономическая интерпретация
25. Реализация методов факторного анализа в современных пакетах прикладных программ. Использование факторного анализа в экономических и социальных исследованиях.
26. Введение в теорию качественных признаков и нечисловой информации
27. Роль и сущность статистики нечисловой информации в экономике, управлении, финансах, социальных науках.
28. Числовые (интервальная, отношений и абсолютная) и нечисловые (номинальная и порядковая) шкалы измерений. Дихотомическая шкала.

**Формы контроля самостоятельной работы обучающихся:** тестирование и расчетное практическое задание.

**Литература по теме:**

1. Мхитарян В. С., Архипова М. Ю., Сиротин В. П. Эконометрика. Учебно-практическое пособие, Место изд.: М., Изд.: Евразийский открытый институт, Год издания: 2012г. // <http://biblioclub.ru/>

2. Ермолаев М. Б., Кадамцева Г. Г., Лапшинов С. Б. Эконометрика. Учебное пособие, Место изд.: Иваново, Изд.: Институт бизнеса, информационных технологий и финансов, Год издания: 2011г.

3. Гмурман, Теория вероятностей и математическая статистика: учеб. пособие для студ. вузов, Место изд.: М., Изд.: Юрайт, Год издания: 2010г.

**Список полезных Интернет-ресурсов:**

[http:// http://biblioclub.ru/](http://http://biblioclub.ru/)

[http:// www.cisstat.org](http://www.cisstat.org)

[http:// www.exponenta.ru](http://www.exponenta.ru)

**Тема 2. Программные средства для обработки данных и системы Data Mining**

**Цель:** заключается в получении обучающимися сведений о принципах разработки и анализа алгоритмов, программ, в соответствии с ОПК-2, с последующим применением в профессиональной сфере и формирование практических навыков по современным методам исследования и информационно-коммуникационными технологиями и новым методам исследования и обработки данных и их применению в самостоятельной деятельности (УК-7).

**Перечень изучаемых элементов содержания**

Программа Excel. Программа Statistica. SPSS. Другие программы. Реляционные базы данных. Параллельные базы данных. Распределённые файловые системы. NoSQL СУБД. Технология Map-Reduce. GOOGLE BIGTABLE. MapReduce. Обычный поиск. Полнотекстовый поиск. Параллельные запросы. Технология поиска и интеграции. Программные средства. ETL процесс по обработке отчётов. Понятие о технологии Data Mining. Реализация в пакетах прикладных программ. Сетевые технологии Data Mining. Примеры применения в социологии и экономике.

**Вопросы для самоподготовки**

1. Понятие о больших данных
2. Программа Excel.
3. Программа Statistica.
4. SPSS.
5. Другие программы.
6. Методы оцифровки
7. Реализация статистических методов в пакетах прикладных программ
8. Формы представления данных: таблицы сопряженности разного вида, кодирование.
9. Методы обработки данных. Навыки работы со статистическими таблицами.
10. Понятие о технологии Data Mining. Реализация в пакетах прикладных программ.
11. Сетевые технологии Data Mining.
12. Примеры применения в социологии и экономике.
13. Реляционные базы данных. Параллельные базы данных.
14. Новые технологии обработки и хранения больших данных
15. Распределённые файловые системы.
16. NoSQL СУБД.
17. Технология Map-Reduce.
18. GOOGLE BIGTABLE.

19. MapReduce.
20. Технологии поиска. Обычный поиск.
21. Полнотекстовый поиск. Параллельные запросы.
22. Интеграция данных из различных источников.
23. Технология поиска и интеграции. Программные средства.
24. ETL процесс по обработке отчётов.

**Формы контроля самостоятельной работы обучающихся:** тестирование и расчетное практическое задание.

**Литература по теме:**

1. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим. Место изд.: М., Изд.: Манн, Иванов и Фербер, Год издания: 2018г. <http://www.ozon.ru/context/detail/id/24323469/>
2. Леонид Черняк. Свежий взгляд на Большие Данные // Открытые системы.СУБД. — 2019. — № 7. — С. 48–51.
3. Леонид Черняк. Что делать с хаосом данных? // Открытые системы.СУБД. — 2013. — № 9. — С. 16–20.
4. Леонид Черняк. Вычисления с акцентом на данные // Открытые системы.СУБД. — 2008. — № 8. — С. 36–39.
5. Леонид Черняк. Смутное время СУБД // Открытые системы.СУБД. — 2012. — № 2. — С. 16–21.

**Список полезных Интернет-ресурсов:**

- <http://www.osp.ru/os/2014/01/13039646/>  
<http://www.ozon.ru/context/detail/id/24323469/>  
<http://www.lookatme.ru/mag/how-to/jobs/202299-big-data> <http://www.osp.ru/os/2011/10/13010990/>

**6. Лабораторный практикум – не предусмотрен**

**7. Практические занятия (семинары) (при наличии)**

<b>Текущий контроль</b>	
<b>Индекс оценочного средства</b>	<b>Название темы/оцениваемого блока (оцениваемых блоков) дисциплины (практики)</b>
<b>1</b>	<b>Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации</b>
Содержание задания для ежедневного/рубежного контрольно- проверочного мероприятия	<i>1.Тестовые задания содержат 15 вопросов теоретического и практического содержания.</i>
Требования к выполнению тестового задания	<i>1.Выполняется письменно. 2.Верное выполнение теоретических и практических тестовых заданий. 3.Врем выполнения тестовых заданий – 30 мин.</i>
Критерии оценки по содержанию и качеству выполнения тестового задания	<i>1.Критерии оценки выполнения тестового задания: «верно»; «неверно». 2.Верное выполнение оценивается в 1 (один) балл. 3.Неверно выполненное отдельное тестовое задание оценивается в 0 баллов.</i>

Методика обработки и форматы представления результатов оценочных процедур	<p>1. При обработке результатов оценочной процедуры используются (ключи, оценочные листы):</p> <ul style="list-style-type: none"> <li>-верное выполнение оценивается в 1 (один) балл;</li> <li>-неверно выполненное отдельное тестовое задание оценивается в 0 баллов;</li> <li>-неявка оценивается в 0 баллов.</li> </ul> <p>2. Результаты оценочной процедуры представляются обучающимся в срок не позднее 1 недели после проведения процедуры текущего контроля и (или) промежуточной аттестации.</p>
<b>2</b>	<b>Тема 2. Программные средства для обработки данных и системы Data Mining</b>
Содержание задания для ежедневного /рубежного контрольно-проверочного мероприятия	1. Тестовые задания содержат 15 вопросов теоретического и практического содержания.
Требования к выполнению тестового задания	<p>1. Выполняется письменно.</p> <p>2. Верное выполнение теоретических и практических тестовых заданий.</p> <p>3. Время выполнения тестовых заданий – 30 мин.</p>
Критерии оценки по содержанию и качеству выполнения тестового задания	<p>1. Критерии оценки выполнения тестового задания: «верно»; «неверно».</p> <p>2. Верное выполнение оценивается в 1 (один) балл.</p> <p>3. Неверно выполненное отдельное тестовое задание оценивается в 0 баллов.</p>
Методика обработки и форматы представления результатов оценочных процедур	<p>1. При обработке результатов оценочной процедуры используются (ключи, оценочные листы):</p> <ul style="list-style-type: none"> <li>-верное выполнение оценивается в 1 (один) балл;</li> <li>-неверно выполненное отдельное тестовое задание оценивается в 0 баллов;</li> <li>-неявка оценивается в 0 баллов.</li> </ul> <p>2. Результаты оценочной процедуры представляются обучающимся в срок не позднее 1 недели после проведения процедуры текущего контроля и (или) промежуточной аттестации.</p>

**Контрольным мероприятием промежуточной аттестации** обучающихся по учебной дисциплине является **зачет**, который проводится в **письменной** форме.

#### **8. Описание материально-технической базы:**

Электронные учебные материалы, используемые преподавателями в образовательном процессе, мультимедийные презентации, банк тестовых заданий и др. представлены на порталах Economist и Web-local.

Для проведения занятий используется следующее оборудование:

- аудиторная доска – 1 шт.;
- мультимедийный проектор – 1 шт.;
- экран – 1 шт.;



– персональные компьютеры (ноутбуки, планшеты) для практических занятий.

#### Описание аудиторий, в которых проводятся занятия

п/п	Фактический адрес учебных кабинетов и объектов	Перечень основного оборудования
1.	ул. Миклухо-Маклая, 6, аудитория 419	мультимедийный проектор, экран, аудиторная доска

## 9. Информационное обеспечение дисциплины

### Перечень информационных технологий. Информационно-справочные системы

В курсе используется следующее специализированное программное обеспечение:

а) программное обеспечение

1. Надстройка Excel для исследования автокорреляционной функции
2. Надстройка Excel для анализа временных рядов методом ARIMA
3. Надстройка Excel для сценарного анализа МонтеКарло
4. Надстройка Excel для ABC-анализа «ABC-Analysis»
5. Система управления проектами Microsoft Project

Ссылки для скачивания файлов:

<http://www.osp.ru/os/2014/01/13039646/>

<http://www.ozon.ru/context/detail/id/24323469/>

<http://www.lookatme.ru/mag/how-to/jobs/202299-big-data>

<http://www.osp.ru/os/2011/10/13010990/>

<http://biblioclub.ru/>

<http://www.cisstat.org>

<http://www.exponenta.ru>

[www.internet-web.ru/](http://www.internet-web.ru/)

информационно-поисковая система Google [www.google.ru](http://www.google.ru);

всемирная свободная Интернет-энциклопедия Wikipedia [ru.wikipedia.org](http://ru.wikipedia.org);

Web-портал компании Microsoft для поддержки программы Excel [office.microsoft.com/ru-ru/access](http://office.microsoft.com/ru-ru/access);

академическая программа Softline Academy Alliance [www.it-academy.ru](http://www.it-academy.ru);

Web-ресурс по информационным технологиям INTERFACE.RU [www.interface.ru](http://www.interface.ru);

Web-портал компании Microsoft в России [www.microsoft.com/ru-ru](http://www.microsoft.com/ru-ru);

Web-портал по продуктам компании Microsoft [msdn.microsoft.com/ru-ru](http://msdn.microsoft.com/ru-ru).

## 10. Учебно-методическое обеспечение дисциплины:

### а) Основная литература.

1. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим. Место изд.: М., Изд.: Манн, Иванов и Фербер, Год издания: 2018г. <http://www.ozon.ru/context/detail/id/24323469/>
2. Мхитарян В. С., Архипова М. Ю., Сиротин В. П. Эконометрика. Учебно-практическое пособие, Место изд.: М., Изд.: Евразийский открытый институт, Год издания: 2019. // <http://biblioclub.ru/>
3. Леонид Черняк. Свежий взгляд на Большие Данные // Открытые системы.СУБД. — 2013. — № 7. — С. 48–51.
4. Гмурман, Теория вероятностей и математическая статистика: учеб. пособие для студ. вузов, Место изд.: М., Изд.: Юрайт, Год издания: 2010г.

### **б) дополнительная литература**

1. Ермолаев М. Б., Кадамцева Г. Г., Лапшинов С. Б. Эконометрика. Учебное пособие, Место изд.: Иваново, Изд.: Институт бизнеса, информационных технологий и финансов, Год издания: 2011г.
2. Леонид Черняк. Что делать с хаосом данных? // Открытые системы.СУБД. — 2013. — № 9. — С. 16–20.
3. Леонид Черняк. Вычисления с акцентом на данные // Открытые системы.СУБД. — 2008. — № 8. — С. 36–39.
4. Леонид Черняк. Смутное время СУБД // Открытые системы.СУБД. — 2012. № 2. — С. 16–21.

### **11. Методические указания для обучающихся по освоению дисциплины (модуля)**

Реализация курса предусматривает интерактивные лекции, практические занятия (семинары) с использованием мультимедийного оборудования, подготовку самостоятельных творческих работ и их последующие презентации, тестирование, проведение групповых дискуссий по тематике курса, современные технологии контроля знаний.

Изучая дисциплину, студент должен прослушать курс лекций, пройти предусмотренное рабочей программой количество семинарских занятий, выполнить задания самостоятельной работы, самостоятельно изучить некоторые темы курса и подтвердить свои знания в ходе контрольных мероприятий.

Работа студента на лекции заключается в уяснении основ дисциплины, кратком конспектировании материала, уточнении вопросов, вызывающих затруднения. Конспект лекций является базовым учебным материалом наряду с учебниками, рекомендованными в основном списке литературы.

Преподавание основной части лекционного материала происходит с использованием средств мультимедиа, которые облегчают восприятие и запоминание материала. Презентации доступны для скачивания с сайта РУДН и могут свободно использоваться студентами в учебных целях.

Студент обязан освоить все темы, предусмотренные учебно-тематическим планом дисциплины. Отдельные темы и вопросы обучения выносятся на самостоятельное изучение. Студент изучает рекомендованную литературу и кратко конспектирует материал, а наиболее сложные вопросы, требующие разъяснения, уточняет во время консультаций. Аналогично следует поступать с разделами курса, которые были пропущены в силу различных обстоятельств.

Для углублённого изучения вопроса студент должен ознакомиться с литературой из дополнительного списка и специализированными сайтами в Интернет. Рекомендуются так же общение студентов на форумах профессиональных сообществ.

Студенты самостоятельно изучают учебную, научную и периодическую литературу. Они имеют возможность обсудить прочитанное с преподавателями дисциплины во время плановых консультаций, с другими студентами на семинарах, а также на лекциях, задавая уточняющие вопросы лектору.

Контроль самостоятельной работы магистров осуществляет ведущий преподаватель. В зависимости от методики преподавания могут быть использованы следующие формы текущего контроля: краткий устный или письменный опрос перед началом занятий, письменное домашнее задание, расчетно-графическая работа и пр.

12. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине (модулю)

Паспорт фонда оценочных средств по дисциплине «Управление процессами на основе больших данных»

Код контролируемой компетенции или ее части	Контролируемый раздел дисциплины	Контролируемая тема дисциплины	ФОСы (формы контроля уровня освоения ООП)									Баллы темы	Баллы раздела	
			Аудиторная работа					Самостоятельная работа						Экзамен/Зачет
			Опрос	Тест	Коллоквиум	Контрольная работа	Дискуссия	Эссе	Выполнение ДЗ	Реферат	Творческий проект	Выполнение КР/КП		
УК-1 УК-2, УК-5, ПКО-15	Раздел 1 Методы многомерного статистического анализа и анализа нечисловой	1. Простейшие статистические характеристики. Приведение к нормальной форме. Оцифровка нечисловых данных. Предмет и содержание раздела «Многомерные статистические методы». Роль и сущность многомерных статистических методов в экономике, управлении, финансах, социальных науках: постановка основных задач, примеры практического использования в социально-экономических исследованиях.	2					5					7	21
		2. Многомерное нормальное распределение как основная модель современных многомерных статистических методов.	2				5						7	
		3. Методы статистического оценивания многомерных параметров и проверки гипотез. Особенности анализа	2					5					7	

		количественных и качественных признаков. Методы шкалирования. Кластерный анализ. Компонентный анализ. Факторный анализ.													
УК-1 УК-2, УК-5, ПКО-15	Раздел 2 Программные средства для обработки данных и системы Data Mining	1. Программа Excel. Программа Statistica. SPSS. Другие программы. Реляционные базы данных. Параллельные базы данных. Распределённые файловые системы. NoSQL СУБД.	2										2	16	
		2. Технология Map-Reduce. GOOGLE BIGTABLE. MapReduce. Обычный поиск. Полнотекстовый поиск.	2			5							7		
		3. Параллельные запросы. Технология поиска и интеграции. Программные средства. ETL процесс по обработке отчётов.													
		4. Понятие о технологии Data Mining. Реализация в пакетах прикладных программ. Сетевые технологии Data Mining. Примеры применения в социологии и экономике..	2					5							7
УК-1 УК-2, УК-5, ПКО-15		Реферат										10		10	
УК-1 УК-2, УК-5, ПКО-15		Рубежная аттестация (контрольная работа)				10								10	
УК-1 УК-2, УК-5, ПКО-15		Тест		10										10	
		Экзамен										20		20	
		ИТОГО	25	10		10	10		15	10			20	100	

**Примерный перечень контрольных заданий/ перечень вопросов к текущему контролю успеваемости и промежуточной аттестации.**

Содержание заданий в составе оценочных средств к текущему контролю успеваемости:

**Примеры тестовых заданий:**

**Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации**

1. Принятый способ представления данных: показатели должны быть:
  - 1) по строкам;
  - 2) по столбцам;
  - 3) по ячейкам;
  - 4) по диагонали.
2. Интервальные данные – это (подчеркните правильные ответы):
  - 1) данные с интервалом;
  - 2) данные об интервалах;
  - 3) количество измерений в каждом интервале;
  - 4) количество интервалов в каждом измерении.
3. Среди ниже приведённых нечисловые данные следующие:
  - 1) баллы;
  - 2) дихотомические;
  - 3) ранги;
  - 4) рейтинги.
4. Среди ниже приведённых нечисловые данные следующие:
  - 1) баллы;
  - 2) дихотомические;
  - 3) ранги;
  - 4) рейтинги.
5. Простейшие статистические характеристики – это:
  - 1) среднее;
  - 2) математическое ожидание;
  - 3) с.к.о.;
  - 4) дисперсия.
6. Приведение к нормальной форме - это:
  - 1) деление на с.к.о.;
  - 2) округление;
  - 3) деление на среднее;
  - 4) деление на константу интегрирования.
7. Какие функции Excel имеют отношение к оцифровке:
  - 1) РАНГ;
  - 2) КОРРЕЛ;
  - 3) СЧЁТЕСЛИ;
  - 4) СУММЕСЛИ.
8. Многомерность в статистике - это:
  - 1) переменных больше одной;
  - 2) переменных больше двух;
  - 3) измерений больше 10;
  - 4) измерений больше 5.
9. Следующие программы являются специализированными статистическими пакетами:
  - 1) EXCEL;
  - 2) SPSS;
  - 3) GRAPHER;
  - 4) STATISTICA.
10. Проверка статистической гипотезы включает в себя:
  - 1) ранжирование;
  - 2) принятие уровня значимости;
  - 3) вычисление эмпирического значения;
  - 4) вычисление критического значения.
11. Кластерный анализ предназначен для:
  - 1) группировки объектов;
  - 2) группировки показателей;
  - 3) ранжирования объектов;
  - 4) ранжирования показателей.
12. Опции кластерного анализа:
  - 1) расстояние между группами;
  - 2) расстояние между показателями;
  - 3) расстояние между объектами;
  - 4) расстояние между телами.
13. Кластерный анализ реализован в программах:
  - 1) EXCEL;
  - 2) AGRAPHER;
  - 3) SPSS;
  - 4) STATISTICA.
14. Снижение размерности это:
  - 1) уменьшение числа измерений;
  - 2) уменьшение числа объектов;
  - 3) уменьшение числа показателей;
  - 4) уменьшение числа знаков.
15. Компонентный анализ реализован в программах:

- 3) EXCEL; 3) SPSS;  
 4) AGRAPHER; 4) STATISTICA.
16. Методы, относящиеся к снижению размерности:  
 1) Факторный анализ; 3) регрессия;  
 2) компонентный анализ; 4) корреляция.
17. Компонентный анализ позволяет:  
 1) сортировать; 3) ранжировать;  
 2) группировать; 4) упорядочивать.
18. Дихотомическая шкала это:  
 1) состоящая из “да” и “нет”; 3) состоящая из двух чисел;  
 2) состоящая из “истина” и “ложь”; 4) состоящая из двух рангов.
19. К нечисловым шкалам относятся:  
 1) номинальная; 3) абсолютная;  
 2) интервалов; 4) ранговая.
20. Существует шкал для описания данных:  
 1) 4; 3) 6;  
 2) 5; 4) 7.
21. Количество наблюдений - это:  
 1) размерность; 3) ширина;  
 2) объём выборки; 4) поверхность выборки.
22. Элементы таблицы сопряжённости называются:  
 1) координаты; 3) скорости;  
 2) длины; 4) частоты.
23. Методы анализа таблиц сопряжённости:  
 1) Критерий Розенбаума; 3) хи-квадрат;  
 2) Критерий Колмогорова-Смирнова; 4) критерий Фишера.
24. В ходе анализа таблицы сопряжённости выполняется:  
 1) проверка на соответствие; 3) проверка на непротиворечивость;  
 2) проверка на монотонность; 4) проверка на значимость.
25. Максимальная размерность таблицы сопряжённости может быть:  
 1) 3; 3) 5;  
 2) 10; 4) какая угодно.
26. Вычисляемое значение критерия хи-квадрат называется:  
 1) Численное значение; 3) реальное значение;  
 2) экспериментальное значение; 4) эмпирическое значение.
27. Вычисляемое значение хи-квадрат сравнивается с:  
 1) критическим значением; 3) предельным значением;  
 2) эталонным значением; 4) граничным значением.
28. То, с чем сравнивается вычисляемое значение хи-квадрат, вычисляется в EXCEL функцией:  
 1) ХИ2РАСП; 3) ХИ2ТЕСТ;  
 2) ХИ2ОБР; 4) ХИ2.
29. К коэффициентам связи относятся:  
 1) коэффициент контингенции; 3) коэффициент ассоциации;  
 2) Коэффициент Чупрова-Крамера; 4) коэффициент коллигации.
30. К разновидности критерия хи-квадрат относятся:  
 1) критерий Вилкоксона; 3) информационный критерий;  
 2) критерий Джонкира; 4) критерий максимального правдоподобия.
31. Выявление вкладов, вносимых каждой клеткой таблицы, называется:  
 1) разбиение хи-квадрат; 3) локализация хи-квадрат;  
 2) анализ хи-квадрат; 4) сортировка хи-квадрат.
32. Лог-линейный анализ - это:

- 1) анализ синтеза таблиц;
- 2) статистический анализ связи таблиц;
- 3) анализ достоверности таблиц;
- 4) анализ разброса таблиц.

## **Тема 2. Программные средства для обработки данных и системы Data Mining**

1. Настоящее и будущее операционных систем Data Mining
2. Методы оцифровки
3. Формы представления данных: таблицы сопряженности разного вида, кодирование.
4. Методы первичной обработки данных. Навыки работы со статистическими таблицами.
5. Понятие о технологии Data Mining.
6. Примеры реализации в пакетах прикладных программ.
7. Сетевые технологии Data Mining.
8. Примеры применения в социологии и экономике.
9. Понятие о базах данных
10. Реляционные базы данных
11. Параллельные базы данных.
12. Новые технологии обработки и хранения больших данных
13. Распределённые файловые системы. NoSQL СУБД.
14. Технология Map-Reduce.
15. Технология GOOGLE BIGTABLE.
16. Технология MapReduce.
17. Технологии поиска
18. Обычный поиск.
19. Полнотекстовый поиск.
20. Параллельные запросы.
21. Интеграция данных из различных источников.
22. Технология поиска и интеграции.
23. ETL процесс по обработке отчётов.

### ***Темы рефератов***

1. Понятие о больших данных
2. Простейшие методы обработки
3. Многомерные статистические методы в экономике, управлении и финансах
4. Реализация статистических методов в пакетах прикладных программ
5. Математические основы многомерных статистических методов
6. Методы множественного корреляционно-регрессионного анализа.
7. Постановка задач классификации
8. Кластерный анализ
9. Использование кластерного анализа
10. Постановка задач снижения размерности
11. Использование компонентного анализа
12. Факторный анализ
13. Использование факторного анализа
14. Введение в теорию качественных признаков и нечисловой информации
15. Методы оцифровки
16. Введение в методы Data Mining
17. Базы данных
18. Новые технологии обработки и хранения больших данных
19. Технологии поиска
20. Интеграция данных из различных источников.

### ***Примерный перечень вопросов к зачету***

### **1. Понятие о больших данных**

Показатели и объекты (измерения). Интервальные данные. Нечисловые данные.

### **2. Простейшие методы обработки**

Простейшие статистические характеристики. Приведение к нормальной форме.

Оцифровка нечисловых данных.

### **3. Многомерные статистические методы в экономике, управлении и финансах**

Предмет и содержание раздела «Многомерные статистические методы». Роль и сущность многомерных статистических методов в экономике, управлении, финансах, социальных науках: постановка основных задач, примеры практического использования в социально-экономических исследованиях.

### **4. Реализация статистических методов в пакетах прикладных программ**

Программа Excel. Программа Statistica. SPSS. Другие программы.

### **5. Математические основы многомерных статистических методов**

Многомерное нормальное распределение как основная модель современных многомерных статистических методов. Практическое применение многомерных методов в финансовых, экономических и социальных исследованиях.

### **6. Методы множественного корреляционно-регрессионного анализа.**

Методы статистического оценивания многомерных параметров и проверки гипотез. Особенности анализа количественных и качественных признаков. Методы шкалирования.

### **7. Постановка задач классификации и кластерный анализ**

Постановка основных прикладных задач классификации многомерных наблюдений. Классификация с обучением и без обучения. Сущность методов классификации. Меры однородности объектов. Расстояния между объектами. Расстояния между кластерами. Реализация методов кластерного анализа в современных пакетах прикладных программ.

### **8. Использование кластерного анализа**

Кластерный анализ финансовой деятельности предприятий. Кластерный анализ мировой демографической статистики. Кластерный анализ социологических опросов. Кластерный анализ результатов аттестации персонала компании. Зависимость выбора метода классификации от цели исследования.

### **9. Постановка задач снижения размерности**

Компонентный анализ. Математическая модель главных компонент. Геометрическая интерпретация главных компонент. Формирование названий главных компонент.

### **10. Использование компонентного анализа**

Экономическая интерпретация главных компонент. Реализация методов компонентного анализа в современных пакетах прикладных программ. Использование компонентного анализа в экономических и социальных исследованиях.

### **11. Факторный анализ**

Линейная модель факторного анализа. Различие предпосылок компонентного и факторного анализа. Экономическая интерпретация

### **12. Использование факторного анализа**

Реализация методов факторного анализа в современных пакетах прикладных программ. Использование факторного анализа в экономических и социальных исследованиях.

### **13. Введение в теорию качественных признаков и нечисловой информации**

Роль и сущность статистики нечисловой информации в экономике, управлении, финансах, социальных науках. Числовые (интервальная, отношений и абсолютная) и нечисловые (номинальная и порядковая) шкалы измерений. Дихотомическая шкала.

### **14. Методы оцифровки**

Формы представления данных: таблицы сопряженности разного вида, кодирование. Методы первичной обработки данных. Навыки работы со статистическими таблицами.



### **15. Введение в методы Data Mining**

Понятие о технологии Data Mining. Реализация в пакетах прикладных программ. Сетевые технологии Data Mining. Примеры применения в социологии и экономике.

### **16. Базы данных**

Реляционные базы данных. Параллельные базы данных.

### **17. Новые технологии обработки и хранения больших данных**

Распределённые файловые системы. NoSQL СУБД. Технология Map-Reduce. GOOGLE BIGTABLE. MapReduce.

### **18. Технологии поиска**

Обычный поиск. Полнотекстовый поиск. Параллельные запросы.

### **19. Интеграция данных из различных источников.**

Технология поиска и интеграции. Программные средства. ETL процесс по обработке отчётов.

### ***Оценивание результатов обучения по учебной дисциплине на промежуточной аттестации***

Ответ обучающегося на зачете оцениваются каждым педагогическим работником по **четырёхбалльной системе**, оценка по учебной дисциплине выставляется в соответствии с Положением о текущем контроле успеваемости и промежуточной аттестации студентов/аспирантов в РГСУ.

#### **Критерии оценки ответа на зачете:**

Оценка «Зачтено» - выставляется при условии, если аспирант показывает хорошие знания изученного учебного материала; самостоятельно, логично и последовательно излагает и интерпретирует материалы учебного курса; полностью раскрывает смысл предлагаемых вопросов; владеет основными терминами и понятиями изученного курса; показывает умение переложить теоретические знания на предполагаемый практический опыт.

Оценка «Не зачтено» - выставляется при наличии серьезных упущений в процессе изложения учебного материала; в случае отсутствия знаний основных понятий и определений курса или присутствии большого количества ошибок при интерпретации основных определений; если аспирант показывает значительные затруднения при ответе на предложенные основные и дополнительные вопросы; при условии отсутствия ответа на основной и дополнительный вопросы.

Материалы для оценки уровня освоения учебного материала дисциплины (оценочные материалы), включающие в себя перечень компетенций с указанием этапов их формирования, описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания, типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы, методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций, разработаны в полном объеме и доступны для обучающихся на странице дисциплины в ТУИС РУДН.

Программа составлена в соответствии с требованиями ОС ВО РУДН.

Доцент кафедры Прикладной экономики

В.П. Корнеенко

\_\_\_\_\_  
должность, название кафедры

\_\_\_\_\_  
подпись

\_\_\_\_\_  
инициалы, фамилия

**РУКОВОДИТЕЛЬ ПРОГРАММЫ**

Зав. кафедрой Прикладной экономики, д.э.н., профессор

  
\_\_\_\_\_

А.А. Чурсин