

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

**ЭКОНОМИЧЕСКИЙ ФАКУЛЬТЕТ
ОТДЕЛЕНИЕ ЭКОНОМИКИ
кафедра Микроэкономики**

УТВЕРЖДЕНА

на заседании кафедры Микроэкономики
Протокол от «01» сентября 2017 г. № 1с

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
(авторский курс)**

Индекс Б1.В.ДВ.1.1 «Введение в большие данные»

по направлению подготовки 38.04.01 Экономика

направленность «Экономика и Финансы»

квалификация Магистр

очная форма обучения

Год набора - 2018

Москва, 2017 г.

Автор—составитель: Жильчиков Алексей Владимирович

Заведующий кафедрой

микроэкономики д.э.н. проф. кафедры микроэкономики Левин М.И.

СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.....	4
2. Объем и место дисциплины в структуре образовательной программы.....	4
3. Содержание и структура дисциплины	5
4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине	5
5. Методические указания для обучающихся по освоению дисциплины	8
6. Учебная литература и ресурсы информационно-телекоммуникационной сети «Интернет», учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине	11
6.1. Основная литература	11
6.2. Дополнительная литература	11
6.3. Учебно-методическое обеспечение самостоятельной работы	11
6.4. Нормативные правовые документы	11
6.5. Интернет-ресурсы	11
6.6. Иные источники	11
7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы	11

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы

1.1. Дисциплина «Введение в большие данные» обеспечивает овладение следующими компетенциями:

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ПК-9	способен анализировать и использовать различные источники информации для проведения экономических расчетов	ПК-9.У	углубление (расширение) сформированной компетенции при условии выбора дисциплины по выбору, элективной дисциплины и/или факультатива

1.2. В результате освоения дисциплины у студентов должны быть сформированы:

ОТФ/ТФ профессиональные действия	Код этапа освоения компетенции	Результаты обучения
ведение аналитической работы в области экономики и финансов	ПК-9.У	на уровне знаний: системы больших данных на уровне умений: применять методы обработки данных на уровне навыков: загрузки данных из разных источников

2. Объем и место дисциплины в структуре ОП ВО

Объем дисциплины

3 ЗЕ, количество академических часов, выделенных на контактную работу с преподавателем - 32 часа, на самостоятельную работу обучающихся - 40 часов;

Место дисциплины в структуре ОП ВО

- Б1.В.ДВ.1.1 «Введение в большие данные», 2 курс, 3 семестр.
- дисциплина реализуется после изучения дисциплин:
 - математический анализ
 - линейная алгебра
 - экономическая информатика
 - иностраный язык
- форма промежуточной аттестации – экзамен.

3. Содержание и структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины, ак. час./час.						Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л	ЛР	ПЗ	КСР		
Очная форма обучения								
Тема 1	Введение в понятие “Большие данные”.	8	2		2		4	О

№ п/п	Наименование тем (разделов)	Объем дисциплины, ак. час./час.						Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л	ЛР	ПЗ	КСР		
Очная форма обучения								
Тема 2	Устройство ОС семейства Linux.	12	2		2		8	ДЗ
Тема 3	Методы обработки текстовых данных	12	2		2		8	ДЗ, КР
Тема 4	Скриптовые языки программирования	16	4		4		8	ДЗ
Тема 5	Реляционные СУБД	14	4		4		6	КР
Тема 6	Преобразование данных	10	2		2		6	К
Промежуточная аттестация		36						экзамен
Всего:		108/81	16/12		16/12		40/30	

Примечание: * – формы текущего контроля успеваемости: контрольная работа (КР), домашнее задание (ДЗ), кейс (К), опрос (О)

Содержание дисциплины

Тема 1. Введение в понятие “Большие данные”.

Тема 2. Устройство ОС семейства Linux.

Тема 3. Методы обработки текстовых данных.

Тема 4. Скриптовые языки программирования.

Язык Perl. Синтаксис языка, основные конструкции (ветвления и циклы). Скаляры, массивы и хэши. Регулярные выражения Perl (PCRE) как основа обработки структурированной текстовой информации. Использование PCRE в других командах ОС Linux.

Тема 5. Реляционные СУБД.

Основные концепции РСУБД. Основы реляционной алгебры. Реляционная полнота языка SQL. Наборы операторов SQL (DDL и DML).

Тема 6. Преобразование данных и доступ к удаленным данным

4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине

4.1. Формы и методы текущего контроля успеваемости.

4.1.1. В ходе реализации дисциплины Б1.В.ДВ.1.1 «Введение в большие данные» используются следующие методы текущего контроля успеваемости обучающихся:

Тема (раздел)	Методы текущего контроля успеваемости
Тема 1	Опрос
Тема 2	Домашнее задание 1
Тема 3	Контрольная работа 1. Домашнее задание 2.
Тема 4	Домашнее задание 3.
Тема 5	Контрольная работа 2.
Тема 6	Кейс

4.1.2. Экзамен проводится с применением следующих методов (средств): в устной форме в виде ответа на теоретические вопросы.

4.2. Материалы текущего контроля успеваемости обучающихся

Типовые оценочные материалы по теме 1

Опрос: Обоснование критериев и результата выбора хостинга

Типовые оценочные материалы по теме 2

Домашнее задание 1. Подготовить сервер для многопользовательской работы с текстовым файлом большого размера:

- создать на сервере необходимое количество пользователей
- создать специальную группу для работы с исходными данными и включить в нее всех созданных пользователей.
- получить файл по заданной ссылке, разархивировать его и настроить общий доступ к нему пользователей созданной группы.

Типовые оценочные материалы по теме 3

Домашнее задание 2. Получить из исходного файла данные за указанный диапазон времени, оставив только колонки данных, необходимых для решения определенной расчетной задачи.

Типовые оценочные материалы по темам 1-3

Контрольная работа 1. Получение трех заданных колонок текстовых данных из строк, содержащих определенное значение и расположенных в строках с 10000 по 12000. Отсеять дубликаты и отсортировать вывод по заданной колонке результата. (Использование конвейра `head ... | tail ... | grep ... | cut ... | sort ...`). Вывод осуществить в файл в своем домашнем каталоге.

Подсчитать и вывести количество строк в результате

Типовые оценочные материалы по теме 4

Домашнее задание 3. Написать на языке perl программу, обрабатывающую исходный файл и получающую результаты идентичные предыдущему заданию.

Типовые оценочные материалы по теме 5

Контрольная работа 2. Создать БД и нужные таблицы для хранения информации о файлах в домашнем каталоге. Разработать процедуру загрузки информации в БД:

- а) Имя файла
- б) Размер
- в) Дата-время создания

Написать SQL запросы для определения:

- а) количества файлов
- б) суммарный размер

Типовые оценочные материалы по теме 6

Кейс: По данным GPS/ГЛОНАСС мониторинга движения транспортных средств оценить эффективность организации маршрутов транспортных средств различными транспортными компаниями.

4.3. Оценочные средства для промежуточной аттестации.

4.3.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы. Показатели и критерии оценивания компетенций с учетом этапа их формирования.

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ПК-9	способен анализировать и использовать различные источники информации для проведения экономических расчетов	ПК-9.У	углубление (расширение) сформированной компетенции при условии выбора дисциплины по выбору, элективной дисциплины и/или факультатива

Этап освоения компетенции	Показатель оценивания	Критерий оценивания
ПК-9.У	способен использовать методы больших данных для проведения экономических расчетов	понимает и объясняет модели, применяемые при анализе больших данных, выбирает и обосновывает применение моделей в зависимости от условий, проводит самостоятельные расчеты

4.3.2. Типовые оценочные средства.

Вопросы для подготовки к экзамену

1. Назначение команды cat. Конвейерные операции.
2. Назначение команды grep. Перенаправление вывода в файл
3. Назначение и принцип работы команды awk
4. Права доступа к файлам и каталогам. Управление правами доступа
5. Копирование, перемещение и удаление файлов и каталогов
6. Основные операции реляционной алгебры
7. Создание БД. Создание таблицы.
8. Оператор SELECT. Получение данных из двух и более таблиц
9. Извлечение информации из БД в файл
10. Загрузка информации из файла в БД
11. Язык Perl. Основы синтаксиса.
12. Язык Perl. Организация циклов и ветвлений.
13. Язык Perl. Чтение файла.
14. Язык Perl. Массивы.
15. Основное назначение регулярных выражений

Шкала оценивания.

Оценка определяется в десятибалльной системе по формуле по формуле:

$$I = 0.1 \cdot O + \frac{\sum_{i=1}^n ДЗ_i}{n} 0.2 + \frac{\sum_{i=1}^n КР_i}{n} 0.2 + 0.25 \cdot K + 0.25 \cdot Э$$

Критерии оценки экзамена

10-балльная шкала	Традиционная шкала	«Зачтено»/ «Не зачтено»	Определение
10	Отлично	Зачтено	Полные, глубокие и систематические знания, знакомство с дополнительной литературой, полный и правильный ответ.

10-бальная шкала	Традиционная шкала	«Зачтено»/ «Не зачтено»	Определение
9	Отлично	Зачтено	Полные, глубокие и систематические знания, полный и правильный ответ.
8	Отлично	Зачтено	Полные и систематические знания, отсутствие существенных неточностей в ответе.
7	Хорошо	Зачтено	Достаточно полные и систематические знания, отсутствие существенных неточностей в ответе.
6	Хорошо	Зачтено	Достаточно полные и систематические знания, с незначительными неточностями в ответе.
5	Удовлетворительно	Зачтено	Знание основного учебного материала в объеме, необходимом для дальнейшей учебы и работы.
4	Удовлетворительно	Зачтено	Знание основного учебного материала в минимальном объеме, необходимом для дальнейшей учебы и работы, имеются погрешности при ответе.
3	Неудовлетворительно	Не зачтено	Допущены существенные ошибки при ответе.
2	Неудовлетворительно	Не зачтено	Имеются пробелы в знаниях по значительной части учебного материала, допущены существенные ошибки при ответе.
1	Неудовлетворительно	Не зачтено	Имеются значительные пробелы в знаниях по значительной части учебного материала, допущены существенные ошибки при ответе.
0	Неудовлетворительно	Не зачтено	Ответ на поставленный вопрос не был сделан

4.4. Методические материалы

Экзамен проводится во время экзаменационной сессии для проверки уровня освоения компетенции. Студент выбирает билет, содержащий один вопрос. Время подготовки к ответу не менее 30 и не более 45 минут. Использование каких-либо дополнительных материалов, а также электронных средств не допускается

5. Методические указания для обучающихся по освоению дисциплины

Домашние задания и примерные варианты контрольных работ для самостоятельной подготовки могут быть опубликованы:

1. В блоге преподавателя на сайте <http://economy.ranepa.ru/prepodavately/>;
2. В группе ВКонтакте;
3. Высланы на общую почту группы

Конкретный способ коммуникации со студентами определяется преподавателем.

Для подготовки к практическим занятиям необходимо ознакомиться с рекомендованной литературой, повторить материал предыдущих практических занятий и лабораторных работ. Самостоятельно разобрать задачи в конце каждой главы. При возникновении вопросов обратиться к преподавателю по электронной почте с указанием конкретной проблемы и (или) прийти к преподавателю на консультацию в установленное время.

Для подготовки письменных домашних заданий необходимо ознакомиться с рекомендованной литературой, повторить материал предыдущих практических занятий и лабораторных работ.

Оформить работу в соответствии с требованиями:

1. В файле с расчетами:
 - Фамилия, Имя, Отчество, номер группы, номер домашнего задания – Section

- Текст задания – Subsection
- Комментарии по заданию - Text
- 2. Имя файла латинскими буквами: DZномер ДЗ-Ivanov(фамилия)
- 3. Тема письма: ЭиФбак-год ДЗномер Фамилия
- 4. Отправить письмо с вложением файла с заданием на электронную почту преподавателя не позднее установленного срока преподавателем.

Для подготовки к контрольным лабораторным работам необходимо ознакомиться с рекомендованной литературой, повторить материал предыдущих практических занятий и лабораторных работ, выполнить типовой вариант контрольной работы. При возникновении вопросов обратиться к преподавателю по электронной почте с указанием конкретной проблемы и (или) прийти к преподавателю на консультацию в установленное время.

Контрольная работа оформляется в соответствии с требованиями:

1. В файле с расчетами:
 - Фамилия, Имя, Отчество, номер группы, номер контрольной работы – Section
 - Задание номер задания– Subsection
 - Комментарии по заданию – Text
 - Комментарии по операторам – (* текст комментария*)
2. Имя файла латинскими буквами: KRномер KP- Ivanov(фамилия)
3. Тема письма: ЭиФбак-год KRномер Фамилия
4. Отправить письмо с вложением файла с заданием на электронную почту преподавателя не позднее установленного срока преподавателем.

В течении курса преподаватель вправе предлагать студентам дополнительные задания повышенной сложности для начисления дополнительных баллов. Правила выполнения данных заданий и начисления баллов объявляются преподавателем индивидуально для каждого задания повышенной сложности.

Самостоятельная работа студентом осуществляется для закрепления изученного материала после практических занятий, для выполнения домашних заданий, для подготовки к контрольным работам, для изучения дополнительных материалов.

№ п/п	Тип занятия	Указания
Тема 1. Введение в понятие “Большие данные”		
1	Л	Возникновение и история отрасли. Определение понятия “большие данные”. Методики анализа больших данных. Вертикальная и горизонтальная масштабируемость.
2	ПЗ	Анализ предложений хостинга виртуальных серверов с точки зрения максимального соответствия задаче обработки на сервере текстового файла с данными размером 4 и более ГБайт. Выработать критериев оценки предложений Составить таблицу предложений, отсортированной по убыванию итоговой оценки.
Тема 2. Устройство ОС семейства Linux.		
3	Л	Система Linux как доступный полнофункциональный набор инструментов для создания систем обработки данных. Основные концепции на примере дистрибутива Debian (Ubuntu). Управление пакетной базой установленной ОС. Базовые команды систем POSIX. Защита от несанкционированного доступа к файлам.

№ п/п	Тип занятия	Указания
4	ПЗ	<p>Установка ОС на виртуальном сервере. Настройка пакетной базы. Знакомство со структурой каталогов ОС Linux, назначение типовых каталогов: /etc, /home, /usr, /var, /proc</p> <p>Работа с командной строкой. Команда и параметры.</p> <p>Команды оценки аппаратной части и доступных ресурсов системы:</p> <p>lshw, lsusb, lspci, ls, du, df</p> <p>Копирование и перемещение файлов между каталогами компьютера и между компьютерами в локальной сети и сети Интернет. Символические и жесткие ссылки.</p> <p>Сходства и различия. Сравнение с ярлыками ОС Windows.</p> <p>Создание и удаление пользователей и групп. Суперпользователь. Предоставление прав доступа на файлы и каталоги пользователям и группам пользователей. Программы su и sudo. Настройка программы sudo.</p> <p>Архивы. Сжатие файлов. Восстановление сжатых файлов.</p> <p>Команды gzip, bzip2, tar.</p> <p>Просмотр содержимого файлов. Команды cat и less.</p> <p>Проработка первой части из 6.1 [2]</p> <p>Выполнить домашнее задание 1.</p>
Тема 3. Методы обработки текстовых данных.		
5	Л	<p>Файлы и потоки. Устройства ввода-вывода как файлы. Принципы конвейерной обработки данных. Текст и структурированный текст как наиболее общий формат хранения данных. Команды обработки текстовых данных в Unix-системах.</p> <p>Выполнить домашнее задание 2.</p> <p>Проработка второй части из 6.1 [2]</p> <p>Подготовка к контрольной работе</p>
6	КР	Контрольная работа 1
Тема 4. Скриптовые языки программирования.		
7	Л	Язык Perl. Синтаксис языка, основные конструкции (ветвления и циклы). Скаляры, массивы и хэши.
8	ПЗ	<p>Скриптовые языки программирования.</p> <p>awk - потоковая обработка структурированных текстовых файлов. Встроенные переменные awk. Получение заданного набора колонок исходного текста.</p> <p>Язык Perl.</p> <p>Проработать главы 1-2 из [3]. Выполнить упражнения (Chapter 2 из [4])</p>
9	Л	Регулярные выражения Perl (PCRE) как основа обработки структурированной текстовой информации. Использование PCRE в других командах ОС Linux.
10	ПЗ	<p>Написание простейшей программы. Изучение поведения переменных разных типов. Область видимости переменных. Ветвления и циклы. Работа с файлами. Регулярные выражения (общие понятия).</p> <p>Проработать главу 3 из [3]. Выполнить упражнения (Chapter 3 из [4])</p> <p>Выполнить домашнее задание 3.</p>
Тема 5. Реляционные СУБД.		
11	Л	Основные концепции РСУБД. Основы реляционной алгебры. Реляционная полнота языка SQL.
12	ПЗ	<p>Консоль SQL. Язык SQL.</p> <p>Операторы определения данных. Создание БД, таблиц и индексов.</p> <p>Изменение структуры БД.</p> <p>Рекомендованная литература: [1]</p>
13	Л	<p>Наборы операторов SQL (DDL и DML).</p> <p>Подготовка к контрольной работе</p> <p>Рекомендованная литература: [1]</p>
14	КР	Контрольная работа 2
Тема 6. Преобразование данных и доступ к удаленным данным		
13	Л	Загрузка и выгрузка текстовой информации в БД.
14	ПЗ	Создание источника данных ODBC для получения удаленных данных для Microsoft Excell.

№ п/п	Тип занятия	Указания
15	Л	ODBC как универсальный метод получения данных. Подключение БД как источника данных к настольным системам
16	ПЗ	Организовать просмотр содержимого домашнего каталога удаленного Linux-сервера средствами MS Excell (используя результаты самостоятельной работы 2). Выполнить итоговую работу.

6. Учебная литература и ресурсы информационно-телекоммуникационной сети «Интернет», включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

6.1. Основная литература.

1. Дейт, К. Дж. Введение в системы баз данных: пер. с англ. / Дейт К. Дж. - 8-е изд. - М.: Вильямс, 2006. - 1328 с. - ISBN 5-8459-0788-8: 624-14.

6.2. Дополнительная литература.

2. Блум, Р., Бреснахэн К. Командная строка Линукс и сценарии оболочки. Библия пользователя. 2-е издание.
3. Рэндал Л. Шварц, Том Феникс, Брайан Де Фой Изучаем Perl. 5-е издание.
<https://www.books.ru/books/izuchaem-perl-5-e-izdanie-666334/>
4. Упражнения по Perl (англ):
<http://www.wiley.com/legacy/compbooks/chapman/perl/exercises.pdf>

6.3. Учебно-методическое обеспечение самостоятельной работы.

Отдельное обеспечение не предусмотрено.

6.4. Нормативные правовые документы.

Не предусмотрены.

6.5. Интернет-ресурсы.

1. Интерпретатор командного языка shell: http://citforum.ru/operating_systems/shell/
2. Онлайн-учебник Perl: <https://ru.perlmaven.com/perl-tutorial>

6.6. Иные источники.

Не предусмотрены.

7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

1. Компьютерный класс
2. Доступ в интернет и локальную сеть Академии
3. Проекционное оборудование
4. Программное обеспечение:
 - Windows/Linux/Mac OS
 - Программа - эмулятор терминала Putty
 - Microsoft Office