

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Цыбиква, Балдыто, Баторович  
Должность: Ректор  
Дата подписания: 11.09.2024 16:43:26  
Уникальный программный ключ:  
056af948c3e48c6f3c571e429957a8ae7b757ae8

**Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Бурятская государственная сельскохозяйственная академия  
имени В.Р. Филиппова»**

**Агрономический факультет**

**СОГЛАСОВАНО**  
Заведующий  
выпускающей кафедрой  
Ландшафтный дизайн и  
экология

\_\_\_\_\_  
уч. ст., уч. зв.

\_\_\_\_\_  
ФИО

\_\_\_\_\_  
подпись

«\_\_» \_\_\_\_\_ 20\_\_ г.

**УТВЕРЖДАЮ**  
Декан агрономического  
факультета

\_\_\_\_\_  
уч. ст., уч. зв.

\_\_\_\_\_  
ФИО

\_\_\_\_\_  
подпись

«\_\_» \_\_\_\_\_ 20\_\_ г.

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ  
дисциплины (модуля)**

**Б1.О.02 Цифровые технологии и анализ данных**

**Направление подготовки  
35.04.01 Лесное дело  
Направленность (профиль)  
Экосистемные услуги на ООПТ  
магистр**

Обеспечивающая преподавание  
дисциплины кафедра  
Разработчик (и)

Информатика и информационные технологии в  
экономике

_____	_____	_____
подпись	уч.ст., уч. зв.	И.О.Фамилия
_____	_____	_____
подпись	уч.ст., уч. зв.	И.О.Фамилия

Внутренние эксперты:  
Председатель методической  
комиссии факультета  
агробизнеса и межкультурных  
коммуникаций

_____	_____	_____
подпись	уч.ст., уч. зв.	И.О.Фамилия

Заведующий методическим  
кабинетом УМУ

_____	_____	_____
подпись		И.О.Фамилия

## ВВЕДЕНИЕ

1. Оценочные материалы по дисциплине (модулю) являются обязательным обособленным приложением к Рабочей программе дисциплины (модуля) и представлены в виде оценочных средств.

2. Оценочные материалы являются составной частью нормативно-методического обеспечения системы оценки качества освоения обучающимися указанной дисциплины (модуля).

3. При помощи оценочных материалов осуществляется контроль и управление процессом формирования обучающимися компетенций, из числа предусмотренных ФГОС ВО в качестве результатов освоения дисциплины (модуля).

4. Оценочные материалы по дисциплине (модулю) включают в себя:

- оценочные средства, применяемые при промежуточной аттестации по итогам изучения дисциплины (модуля).

- оценочные средства, применяемые в рамках индивидуализации выполнения, контроля фиксированных видов ВАРО;

- оценочные средства, применяемые для текущего контроля;

5. Разработчиками оценочных материалов по дисциплине (модулю) являются преподаватели кафедры, обеспечивающей изучение обучающимися дисциплины (модуля) в Академии. Содержательной основой для разработки оценочных материалов является Рабочая программа дисциплины (модуля).

**1. ОЖИДАЕМЫЕ РЕЗУЛЬТАТЫ ИЗУЧЕНИЯ**  
**учебной дисциплины (модуля), персональный уровень достижения которых проверяется**  
**с использованием представленных в п. 3 оценочных материалов**

Компетенции, в формировании которых задействована дисциплина		Код и наименование индикатора достижений компетенции	Компоненты компетенций, формируемые в рамках данной дисциплины (как ожидаемый результат ее освоения)		
код	наименование		знать и понимать	уметь делать (действовать)	владеть навыками (иметь навыки)
1		2	3	4	5
<b>Общепрофессиональные компетенции</b>					
ОПК-1.	Способен анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	ИД 1 ОПК 1 Обосновывает выбор технологических приемов в профессиональной деятельности, опираясь на анализ достижений науки и производства.	Знает, как обосновать выбор технологических приемов в профессиональной деятельности, опираясь на анализ достижений науки и производства	Умеет обосновывать выбор технологических приемов в профессиональной деятельности, опираясь на анализ достижений науки и производства	Владеет навыками обоснования выбора технологических приемов в профессиональной деятельности, опираясь на анализ достижений науки и производства
		ИД 2 ОПК 1 Выявляет перспективные направления повышения эффективности технологических приемов в профессиональной сфере.	Знает перспективные направления повышения эффективности технологических приемов в профессиональной сфере	Применяет перспективные направления повышения эффективности технологических приемов в профессиональной сфере	Владеет технологическими приемами в профессиональной сфере

**2. РЕЕСТР  
элементов оценочных материалов по дисциплине (модулю)**

Группа оценочных средств	Оценочное средство или его элемент
	Наименование
1	2
<b>1. Средства для промежуточной аттестации по итогам изучения дисциплины</b>	Перечень экзаменационных вопросов
	Экзаменационная программа по учебной дисциплине
	Пример экзаменационного билета
	Плановая процедура проведения экзамена
	Критерии оценки к экзамену
<b>2. Средства для индивидуализации выполнения, контроля фиксированных видов (ВАРО)</b>	Не предусмотрены учебным планом
<b>3. Средства для текущего контроля</b>	Комплект заданий для лабораторных работ
	Критерии оценивания лабораторных работ
	Шкала оценивания
	Перечень групповых заданий
	Критерии оценивания групповых заданий
	Шкала оценивания
	Комплект заданий для самостоятельной работы обучающихся
	Критерии оценивания самостоятельной работы
	Шкала оценивания
	Тестовые задания
	Критерии оценивания тестовых заданий
Шкала оценивания	

### 3. Описание показателей, критериев и шкал оценивания компетенций в рамках дисциплины (модуля)

Индекс и название компетенции	Код индикатора достижений компетенции	Индикаторы компетенции и	Показатель оценивания – знания, умения, навыки (владения)	Уровни сформированности компетенций				Формы и средства контроля формирования компетенций
				компетенция не сформирована	минимальный	средний	высокий	
				Оценки сформированности компетенций				
				2	3	4	5	
				Оценка «неудовлетворительно»	Оценка «удовлетворительно»	Оценка «хорошо»	Оценка «отлично»	
				Характеристика сформированности компетенции				
1	2	3	4	5	6	7	8	9
Критерии оценивания								
ОПК-1 Способен анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности;	ИД 1 опк 1 Обосновывает выбор технологических приемов в профессиональной деятельности, опираясь на анализ достижений науки и производства.	Полнота знаний	Обучающийся знает, как анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	Не знает, как анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	Не в полной мере знает, как анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	Хорошо знает, как анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	В полной мере знает, как анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	Перечень вопросов к экзамену, комплект заданий для лабораторных работ, перечень групповых заданий, комплект заданий для самостоятельной работы обучающихся, тестовые задания
		Наличие умений	Обучающийся умеет анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	Не умеет анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	Не в полной мере умеет анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	Хорошо умеет анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	В полной мере умеет анализировать современные проблемы науки и производства, решать сложные (нестандартные) задачи в профессиональной деятельности	

		Наличие навыков (владение опытом)	Обучающийся владеет навыками анализа современных проблем науки и производства, решения сложных (нестандартных) задач в профессиональной деятельности	Не владеет навыками анализа современных проблем науки и производства, решения сложных (нестандартных) задач в профессиональной деятельности	Не в полной мере владеет навыками анализа современных проблем науки и производства, решения сложных (нестандартных) задач в профессиональной деятельности	Хорошо владеет навыками анализа современных проблем науки и производства, решения сложных (нестандартных) задач в профессиональной деятельности	В полной мере владеет навыками анализа современных проблем науки и производства, решения сложных (нестандартных) задач в профессиональной деятельности	
ИД 2 <sub>ОПК-1</sub> Выявляет перспективные направления повышения эффективности технологических приемов в профессиональной сфере	Полнота знаний	Знает перспективные направления повышения эффективности технологических приемов в профессиональной сфере	Не знает перспективные направления повышения эффективности технологических приемов в профессиональной сфере	Не в полной мере знает перспективные направления повышения эффективности технологических приемов в профессиональной сфере	Хорошо знает перспективные направления повышения эффективности технологических приемов в профессиональной сфере	В полной мере знает перспективные направления повышения эффективности технологических приемов в профессиональной сфере		
	Наличие умений	Применяет перспективные направления повышения эффективности технологических приемов в профессиональной сфере	Не применяет перспективные направления повышения эффективности технологических приемов в профессиональной сфере	Не в полной мере применяет перспективные направления повышения эффективности технологических приемов в профессиональной сфере	Хорошо применяет перспективные направления повышения эффективности технологических приемов в профессиональной сфере	В полной мере применяет перспективные направления повышения эффективности технологических приемов в профессиональной сфере		
	Наличие навыков (владение опытом)	Владеет технологическими приемами в профессиональной сфере	Не владеет технологическими приемами в профессиональной сфере	Не в полной мере владеет технологическими приемами в профессиональной сфере	Хорошо владеет технологическими приемами в профессиональной сфере	В полной мере владеет технологическими приемами в профессиональной сфере		

**4. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения основной профессиональной образовательной программы**

**4.1. Типовые контрольные задания, необходимые для оценки знаний, умений, навыков**

**4.1.1. Средства для промежуточной аттестации по итогам изучения дисциплины**

<b>Нормативная база</b> <b>проведения промежуточной аттестации обучающихся по результатам изучения дисциплины:</b> Б1.О.02 Цифровые технологии и анализ данных	
1) действующее «Положение о текущем контроле успеваемости и промежуточной аттестации обучающихся ФГБОУ ВО Бурятская ГСХА»	
<b>Основные характеристики</b> <b>промежуточной аттестации обучающихся по итогам изучения дисциплины (модуля)</b>	
1	2
<b>Цель промежуточной аттестации -</b>	установление уровня достижения каждым обучающимся целей обучения по данной дисциплине, изложенных в п.2.2 настоящей программы
<b>Форма промежуточной аттестации -</b>	экзамен
<b>Место экзамена в графике учебного процесса:</b>	1) подготовка к экзамену и сдача экзамена осуществляется за счёт учебного времени (трудоемкости), отведённого на экзаменационную сессию для обучающихся, сроки которой устанавливаются приказом по академии 2) дата, время и место проведения экзамена определяется графиком сдачи экзаменов, утверждаемым деканом факультета (директором института)
<b>Форма экзамена -</b>	<i>устный</i>
<b>Процедура проведения экзамена -</b>	представлена в оценочных материалах по дисциплине
<b>Экзаменационная программа по учебной дисциплине:</b>	1) представлена в оценочных материалах по дисциплине 2) охватывает разделы (в соответствии с п. 4.1 настоящего документа)
<b>Методические материалы, определяющие процедуры оценивания знаний, умений, навыков:</b>	представлены в оценочных материалах по дисциплине

**Перечень вопросов к зачету**

1. Определите сущность понятия «большие данные» ОПК-1
2. Опишите методики анализа больших данных. ОПК-1
3. Процесс аналитики анализа больших данных. ОПК-1
4. Дайте характеристику Big Data на мировом рынке. ОПК-1
5. Охарактеризуйте Big Data в России. ОПК-1
6. Определите понятие Data Mining. ОПК-1
7. Вопросы безопасности больших данных. ОПК-1
8. В чем состоит когнитивный анализ данных. ОПК-1
9. Какие модели данных вы знаете? ОПК-1
10. Основные описательные статистики. ОПК-1
11. Определите различия между параметрическими, непараметрическими и номинальными методами. ОПК-1
12. Опишите основную идею корреляционного анализа. ОПК-1
13. Регрессионный анализ. ОПК-1
14. Основная идея дисперсионного анализа. ОПК-1
15. Сущность кластерного анализа. ОПК-1
16. Дискриминантный анализ: модель и общая процедура выполнения. ОПК-1
17. Цели факторного анализа. ОПК-1
18. Программные средства анализа данных: Statistica, SPSS, Excel; их преимущества и недостатки. ОПК-1
19. Преимущества работа с данными в программе R-Studio. ОПК-1
20. Представление исходных данных в программе R-Studio. ОПК-1
21. Выполнение анализа данных в R-Studio. ОПК-1
22. Облачные платформы для анализа больших данных. ОПК-1

**Экзаменационная программа по учебной дисциплине**

Разработана на основе положения СТО СМК – 7.6.П-4.0-2019 Положение о текущем контроле успеваемости и промежуточной аттестации обучающихся ФГБОУ ВО Бурятская ГСХА.

Экзаменационные билеты оформляются по следующей форме (образец):

федеральное государственное бюджетное образовательное учреждение высшего образования «Бурятская государственная сельскохозяйственная академия имени В.Р. Филиппова»

Заведующий кафедрой \_\_\_\_\_ / \_\_\_\_\_  
(наименование кафедры) (подпись) (ФИО)

Дисциплина Компьютерное проектирование в ландшафтном дизайне  
Экзаменационный билет №1

Вопросы:

- 1.
- 2.
- 3.

### Плановая процедура проведения экзамена

Экзамен проводится в форме устного опроса по билетам (вопросам) с предварительной подготовкой. Экзаменатор вправе задавать вопросы сверх билета, а также, помимо теоретических вопросов.

Экзаменационные билеты (вопросы) утверждаются на заседании кафедры и подписываются заведующим кафедрой не позднее, чем за две недели до начала экзаменационной сессии. В билете должно содержаться не более трех вопросов. Комплект экзаменационных билетов по дисциплине должен содержать не менее 20 билетов.

### 5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

#### 5.1. Критерии оценки к экзамену

*Оценка «отлично» (86-100 баллов)* ставится обучающемуся, обнаружившему систематические и глубокие знания учебно-программного материала, умения свободно выполнять задания, предусмотренные программой в типовой ситуации (с ограничением времени) и в нетиповой ситуации, знакомство с основной и дополнительной литературой, усвоение взаимосвязи основных понятий дисциплины в их значении приобретаемой специальности и проявившему творческие способности и самостоятельность в приобретении знаний. Студент исчерпывающим образом ответил на вопросы экзаменационного билета. Задача решена правильно, студент способен обосновать выбранный способ и пояснить ход решения задачи.

*Оценка «хорошо» (71-85 баллов)* ставится обучающемуся, обнаружившему полное знание учебно-программного материала, успешное выполнение заданий, предусмотренных программой в типовой ситуации (с ограничением времени), усвоение материалов основной литературы, рекомендованной в программе, способность к самостоятельному пополнению и обновлению знаний в ходе дальнейшей работы над литературой и в профессиональной деятельности. При ответе на вопросы экзаменационного билета студентом допущены несущественные ошибки. Задача решена правильно или ее решение содержало несущественную ошибку, исправленную при наводящем вопросе экзаменатора.

*Оценка «удовлетворительно» (56-70 баллов)* ставится обучающемуся, обнаружившему знание основного учебно-программного материала в объеме, достаточном для дальнейшей учебы и предстоящей работы по специальности, знакомство с основной литературой, рекомендованной программой, умение выполнять задания, предусмотренные программой. При ответе на экзаменационные вопросы и при выполнении экзаменационных заданий обучающийся допускает погрешности, но обладает необходимыми знаниями для устранения ошибок под руководством преподавателя. Решение задачи содержит ошибку, исправленную при наводящем вопросе экзаменатора.

*Оценка «неудовлетворительно» (менее 56 баллов)* ставится обучающемуся, обнаружившему пробелы в знаниях основного учебно-программного материала, допустившему принципиальные ошибки в выполнении предусмотренных программой заданий, слабые побуждения к самостоятельной работе над рекомендованной основной литературой. Оценка «неудовлетворительно» ставится



обучающимся, которые не могут продолжить обучение или приступить к профессиональной деятельности по окончании академии без дополнительных занятий по соответствующей дисциплине.

## **6. Оценочные материалы для организации текущего контроля успеваемости обучающихся**

Форма, система оценивания, порядок проведения и организация текущего контроля успеваемости обучающихся устанавливаются Положением об организации текущего контроля успеваемости обучающихся.

### **Комплект заданий для лабораторных работ**

Защита лабораторной работы осуществляется путем собеседования студента с преподавателем. При собеседовании студент представляет на проверку отчет по лабораторной работе.

#### **Раздел 1. Введение в анализ больших данных.**

Тема: Поиск источников информации в сети Интернет для Big Data

Содержание темы. Сбор и хранение больших данных. Поиск источников информации в сети Интернет: открытые и закрытые источники данных. Портал открытых данных РФ. Сохранение данных в программе Excel. Преобразование и первичная обработка данных.

#### **Раздел 2. Методы анализа данных**

Тема: Методы обработки и анализа больших данных.

Содержание темы. Представление исходных данных в программе R-Studio (векторы, массивы, матрицы, списки, таблицы). Статистическая обработка данных в программах Excel R-Studio: подсчет описательных статистик, графическое представление данных. Группировка данных, обнаружение значимых корреляций, зависимостей и тенденций в результате анализа имеющейся информации, выявления отношений между данными различного типа. Применение различных методов выделения, извлечения и группировки данных, которые позволяют выявить систематизированные структуры данных и вывести из них правила для принятия решений и прогнозирования их последствий (регрессионный, дисперсионный, кластерный, дискриминантный, факторный анализы).

Тема: Визуализация исходной информации и аналитических данных.

Содержание темы. Возможности графического представления информации в программе R-Studio: графические функции отображения одномерных и многомерных данных, графический вывод с использованием графических параметров.

#### **Критерии оценивания:**

- правильность выполнения задания на лабораторную работу;
- степень усвоения теоретического материала по теме лабораторной работы;
- способность продемонстрировать преподавателю навыки работы в инструментальной программной среде, а также применить их к решению типовых задач;
- качество подготовки отчета по лабораторной работе;
- правильность и полнота ответов на вопросы преподавателя при защите работы.

#### **Шкала оценивания**

Баллы для учета в рейтинге (оценка)	Степень удовлетворения критериям
86-100 баллов «отлично»	Выполнены все задания лабораторной работы, обучающийся четко и без ошибок ответил на все контрольные вопросы
71-85 баллов «хорошо»	Выполнены все задания лабораторной работы; обучающийся ответил на все контрольные вопросы с замечаниями
56-70 баллов «удовлетворительно»	Выполнены все задания лабораторной работы с замечаниями; обучающийся ответил на все контрольные вопросы с замечаниями
менее 56 баллов «неудовлетворительно»	Обучающийся не выполнил или выполнил неправильно задания лабораторной работы; обучающийся ответил на контрольные вопросы с ошибками или не ответил на контрольные вопросы

### **Перечень групповых заданий Работа в малых группах**

Группа обучающихся делится на несколько малых групп. Количество групп определяется числом творческих заданий, которые будут обсуждаться в процессе занятия. Малые группы формируются либо по желанию, либо по родственной тематике для обсуждения. Каждая малая группа составляет презентацию по своей теме для дальнейшей защиты.

**Тема: Обзор источников информации для Big Data**

1. Открытые источники информации: статистические сборники, опубликованные отчеты и результаты исследований. Портал открытых данных РФ
2. Закрытые источники данных. Доступ к закрытой информации. Вопросы безопасности данных.

**Тема: Облачные платформы для анализа больших данных**

1. Аналитика больших данных. Преимущества облаков для аналитики больших данных
3. Частные, публичные и гибридные облака для аналитики больших данных
4. Основные сферы применения больших данных для принятия решений. Обзор тенденций и проблем больших данных
5. Мотивирующие примеры применения Big Data в бизнесе

**Критерии оценивания групповых заданий**

- актуальность темы;
- соответствие содержания работы выбранной тематике;
- соответствие содержания и оформления работы установленным требованиям;
- обоснованность результатов и выводов, оригинальность идеи;
- новизна полученных данных;
- личный вклад обучающихся;
- возможности практического использования полученных данных;
- оформление презентации.

**Шкала оценивания**

Баллы для учета в рейтинге (оценка)	Степень удовлетворения критериям
86-100 баллов «отлично»	Работа демонстрирует точное понимание задания. Все материалы имеют непосредственное отношение к теме; источники цитируются правильно. Результаты работы представлены четко и логично, информация точна и отредактирована. Работа отличается яркой индивидуальностью и выражает точку зрения обучающегося. Соблюден единый стиль оформления презентации
71-85 баллов «хорошо»	Помимо материалов, имеющих непосредственное отношение к теме, включаются некоторые материалы, не имеющие отношение к ней; используется ограниченное количество источников. Не вся информация взята из достоверных источников; часть информации неточна или не имеет прямого отношения к теме. Недостаточно выражена собственная позиция и оценка информации. Соблюден единый стиль оформления презентации. Допускается незначительная перегрузка слайда информацией.
56-70 баллов «удовлетворительно»	Часть материалов не имеет непосредственного отношения к теме, используется 2-3 источника. Делается слабая попытка проанализировать информацию. Материал логически не выстроен и подан внешне непривлекательно, не дается четкого ответа на поставленные вопросы. Нет критического взгляда на проблему. Единый стиль оформления нарушен.
менее 56 баллов «неудовлетворительно»	Больше половины материалов не имеет непосредственного отношения к теме, используется один источник. Не делается попытка проанализировать информацию. Материал логически не выстроен и подан внешне непривлекательно, не дается ответа на поставленные вопросы. Презентация не представлена.

**Комплект заданий  
для самостоятельной работы обучающихся**

Изучить теоретический материал в соответствии с темой работы

**Раздел 1. Введение в анализ больших данных.**

1. Определите сущность понятия «большие данные».
2. Опишите методики анализа больших данных.
3. Процесс аналитики анализа больших данных.
4. Дайте характеристику Big Data на мировом рынке.
5. Охарактеризуйте Big Data в России.
6. Определите понятие Data Mining.
7. Вопросы безопасности больших данных.
8. В чем состоит когнитивный анализ данных.
9. Какие модели данных вы знаете?

**Раздел 2. Методы анализа данных**

1. Основные описательные статистики.

2. Определите различия между параметрическими, непараметрическими и номинальными методами.
3. Опишите основную идею корреляционного анализа.
4. Регрессионный анализ.
5. Основная идея дисперсионного анализа.
6. Сущность кластерного анализа.
7. Дискриминантный анализ: модель и общая процедура выполнения.
8. Цели факторного анализа.
9. Программные средства анализа данных: Statistica, SPSS, Excel; их преимущества и недостатки. 19. Преимущества работа с данными в программе R-Studio.
10. Представление исходных данных в программе R-Studio.
11. Выполнение анализа данных в R-Studio.

#### Критерии оценивания самостоятельной работы:

- правильность выполнения задания;
- степень усвоения теоретического материала по теме;
- способность продемонстрировать преподавателю навыки работы в инструментальной программной среде, а также применить их к решению типовых задач;
- качество подготовки отчета по самостоятельной работе;
- правильность и полнота ответов на вопросы преподавателя при защите работы.

#### Шкала оценивания

Баллы для учета в рейтинге (оценка)	Степень удовлетворения критериям
86-100 баллов «отлично»	Выполнены все задания самостоятельной работы, обучающийся четко и без ошибок ответил на все контрольные вопросы
71-85 баллов «хорошо»	Выполнены все задания самостоятельной работы; обучающийся ответил на все контрольные вопросы с замечаниями
56-70 баллов «удовлетворительно»	Выполнены все задания самостоятельной работы с замечаниями; обучающийся ответил на все контрольные вопросы с замечаниями
менее 56 баллов «неудовлетворительно»	Обучающийся не выполнил или выполнил неправильно задания самостоятельной работы; обучающийся ответил на контрольные вопросы с ошибками или не ответил на контрольные вопросы

#### Тестовые задания

1. большинство данных в мире в 2011 году содержалось:
  - 1) в цифровом виде
  - 2) в аналоговом виде
2. В каком веке произошёл перевес объёмов накопленных человечеством данных в сторону цифровых?  
**Ответ: 20**
3. Объём накопленных человечеством цифровых данных на 2012 год измеряется:
  - 1) петабайтами
  - 2) зеттабайтами
  - 3) эксабайтами
  - 4) йоттабайтами
4. Сколько Петабайт в Зеттабайте? Укажите число.  
**Ответ: 1024**
5. укажите фактор, способствовавший появлению тренда больших данных
  - 1) маркетинговые кампании крупных корпораций
  - 2) снижение издержек на хранение данных
  - 3) появление новых технологий обработки потоковых данных
  - 4) выпуск баз данных с обработкой данных в памяти
6. Какие вероятные разочарования тренда больших данных?
  - 1) из-за угрозы безопасности личной жизни (privacy) граждан будут усложнены процедуры сбора данных, что приведёт к падению ценности больших данных
  - 2) из-за угрозы безопасности личной жизни (privacy) граждан будут упрощены процедуры сбора данных, что приведёт к падению ценности больших данных

3) нет

7. Отметьте значимые события, повлиявшие на формирование тренда больших данных:

- 1) **разработка Hadoop**
- 2) **изобретение принципа MapReduce**
- 3) разработка языка Python
- 4) победа Deepblue в матче с Г.Каспаровым

8. Выберите верный ответ

- 1) большие данные – это обработка или хранение более 1 Тб информации
- 2) **проблема больших данных – это такая проблема, когда при существующих технологиях хранения и обработки существенная обработка данных затруднена или невозможна**
- 3) большие данные – это огромная PR-акция крупных вендоров и не более того
- 4) большие данные – это явление, когда цифровые данные наиболее полно представляют изучаемый объект

9. Выберите неверный ответ:

- 1) **большие данные – это данные объёма свыше 1 Тб**
- 2) проблема больших данных – это проблема, когда при существующих технологиях хранения и обработки существенная обработка данных затруднена или невозможна
- 3) большие данные – это тренд в области ИТ, подогреваемый маркетинговыми кампаниями крупных вендоров
- 4) большие данные как правило не структурированы

10. Отметьте те из вариантов, в которых данные структурированы:

- MS Word
- 1) данные о продажах компании, представленные в виде помесечных отчётов в формате MS Word
  - 2) **таблица с ежедневными показаниями температуры помещения за год в файле формата csv**
  - 3) текст педагогической поэмы А.С. Макаренко, представленный в формате PDF
  - 4) библиотека фильмов, представленных в формате mp4 на одном жестком диске

11. Перечислите четыре основных характеристики Big Data:

- 1) Virtualization, Volume, Variability, Velocity
- 2) **Variety, Velocity, Volume, Value**
- 3) Verification, Volume, Velocity, Visualization
- 4) Video, Value, Variety, Volume

12. Выберите неверное высказывание:

- 1) **большие объёмы данных приводят к слабой их структуризации, поэтому появляется такое разнообразие данных**
- 2) увеличившаяся производительность телекоммуникационных каналов привела к росту объёмов передаваемой информации
- 3) удешевление систем хранения на единицу информации привело к росту рынка больших данных

13. Отметьте неверное понимание Variety в контексте характеристик Big Data:

- 1) **высокая скорость генерирования данных**
- 2) **разные типы данных в колонках таблиц реляционных СУБД**
- 3) **разнообразие отраслей, являющихся источниками данных**
- 4) разнообразие типов данных, включающих в себя структурированные, полуструктурированные и неструктурированные

14. Принцип MapReduce состоит в том, чтобы

- 1) **производить вычисления на узлах, где информация изначально была сохранена**
- 2) **использовать вычислительные мощности систем хранения**
- 3) использовать функциональное программирование для решения задач массивно-параллельной обработки

15. Выберите одно неверное высказывание про MapReduce:

- 1) интерфейс для массово-параллельной обработки данных, где вычисления производятся на узлах, где информация изначально была сохранена
- 2) **MapReduce – это две операции: распределения и сборки данных**

3) **MapReduce был придуман разработчиками Hadoop**

4) MapReduce был анонсирован разработчиками Google

16. Во сколько раз теоретически вырастет производительность при подсчёте числа слов в тексте при работе MapReduce при переходе от одного узла к двум? (Введите число.)

**Ответ: 2**

17. Какие из следующих технологий СУБД не используют принцип MapReduce

- 1) Hadoop
- 2) Cassandra
- 3) HDInsight
- 4) **Redis**

18. Какие СУБД полностью полагаются на оперативную память при хранении информации:

- 1) **Oracle Exalytics**
- 2) **SAP HANA**
- 3) BigTable
- 4) HBase

19. В чём преимущество колоночно-ориентированных СУБД?

СУБД

- 1) они позволяют выполнять более сложные SQL-запросы по сравнению с реляционными
- 2) **они позволяют динамически дополнять содержание записей новыми полями**
- 3) они имеют более гибкие возможности аналитики
- 4) они позволяют эффективно делать межколоночные сравнения

20. Для чего аналитику необходима "песочница"?

1) **для высокопроизводительной аналитики за счёт использования оперативной памяти и inDB операций**

- 2) для хранения всех полученных от заказчика данных
- 3) для построения отчётов о результатах анализа
- 4) **для снижения затрат, связанных с репликацией данных**

21. Какие из следующих средств разумно использовать для анализа данных, представленных единственным csv-файлом размера более 100Гб:

- 1) **Hadoop**
- 2) Data Warehouse
- 3) **"Песочница"**
- 4) Python

22. Выберите верное утверждение:

- 1) Data Warehouse создаются для проверки гипотез при анализе больших данных
- 2) **"Песочница" используется для снижения нагрузки на основной Data Warehouse**
- 3) каждый Data Warehouse должен содержать "песочницу"
- 4) "Песочница" необходима для любого процесса аналитики

23. Ниже приведена последовательность этапов проекта аналитики в соответствии с CRISP-DM, укажите первый этап.

- 1) моделирование (Modeling)
- 2) внедрение (Deployment)
- 3) подготовка данных (Data Preparation)
- 4) **понимание бизнеса (Business understanding)**
- 5) оценка (Evaluation)
- 6) понимание данных (Data Understanding)

24. На каком из этапов процесса CRISP-DM происходит проверка гипотез?

- 1) понимание бизнеса (Business understanding)
- 2) понимание данных (Data Understanding)
- 3) **моделирование (Modeling)**
- 4) оценка (Evaluation)

25. Вы являетесь владельцем и аналитиком в компании из 10 человек, в которой требуется проанализировать продажи за 1 год (1 млн. продаж). Какие из этапов CRISP-DM можно опустить:

- 1) понимание бизнеса (Business understanding)
- 2) подготовка данных (Data Preparation)
- 3) моделирование (Modeling)
- 4) оценка (Evaluation)

26. Пример благоразумного использования Hadoop

- 1) анализ 10 Гб данных
- 2) ежедневное сохранение данных температуры, поступающих со всех городов России (по одному показанию на город, всего городов 1100 шт)
- 3) **посекундное сохранение данных температуры, поступающих со всех городов России (по одному показанию на город, всего городов 1100 шт)**
- 4) построение графика пульса пациента в реальном времени

27. Начиная с каких размеров данных обоснованно применение кластера Hadoop для хранения данных?

- 1) 100Гб
- 2) 1Тб
- 3) **100Тб**
- 4) 1Пб

28. Hadoop – это:

- 1) **набор утилит, и программный каркас для выполнения распределённых программ, работающих на кластерах**
- 2) распределённая СУБД, позволяющая обрабатывать большие данные
- 3) язык выполнения заданий в парадигме MapReduce
- 4) распределённая файловая система, предназначенная для хранения файлов большого объёма

#### Критерии оценивания тестовых заданий

– отношение правильно выполненных заданий к общему их количеству

#### Шкала оценивания

Баллы для учета в рейтинге (оценка)	Степень удовлетворения критериям
86-100 баллов «отлично»	выполнено от 86% до 100% теста
71-85 баллов «хорошо»	выполнено от 72% до 85% теста
56-70 баллов «удовлетворительно»	выполнено от 56% до 71% теста
0-55 баллов «неудовлетворительно»	выполнено менее 56% теста