

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Антипова Наталья Викторовна
Должность: и.о. директора филиала
Дата подписания: 26.06.2026 18:02:46
Уникальный программный ключ:
fae5412acb1bf810dc69e6bc004ac45622b84b3a

Приложение 6
к основной профессиональной образовательной программе
по направлению подготовки 38.04.01 Экономика
направленность (профиль) программы «Бизнес-аналитика в
экономике и управлении»

**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 № 3 от 23 октября 2025 г.
Председатель совета
Н.В. Антипова



**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ
по дисциплине
Б1.В.02 Программные средства бизнес-аналитики**

Направление подготовки 38.04.01 Экономика

Направленность (профиль) программы Бизнес-аналитика в экономике и управлении

Уровень высшего образования Магистратура

Год начала подготовки 2026

Улан-Батор – 2025 г.

Составители:

к.э.н., доцент, доцент кафедры статистики РЭУ им.
Г.В. Плеханова

И.П. Комарова

Оценочные материалы одобрены на заседании Совета Улан-Баторского филиала РЭУ им. Г.В. Плеханова
протокол № 3 от «23» октября 2025 г.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

по дисциплине

«Программные средства бизнес-аналитики»

ПЕРЕЧЕНЬ КОМПЕТЕНЦИЙ С УКАЗАНИЕМ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ И ЭТАПОВ ИХ ФОРМИРОВАНИЯ ПО ДИСЦИПЛИНЕ

Формируемые компетенции (код и наименование компетенции)	Индикаторы достижения компетенций (код и наименование индикатора)	Результаты обучения (знания, умения)	Наименование контролируемых разделов и тем
<i>ПК-1 Способен обосновывать подходы, используемые в бизнес-анализе</i>	<i>ПК-1.2 Определяет подход к работе с информацией бизнес-анализа</i>	ПК-1.2. З-1. Знает источники информации и аналитические методы подготовки обоснования выбора управленческих решений ПК-1.2. У-1. Умеет работать с базами данных, интернет-ресурсами, распространенными программами для бизнеса, подготовки сводных, аналитических отчетов для управления компанией с учетом поставленной цели и задач	Тема 1. Первичная обработка данных. Графические средства анализа Тема 2. Методы классификации. «Случайный лес» Тема 3. Компонентное моделирование. Факторный анализ Тема 4. Ассоциативное моделирование
<i>ПК-2 Способен руководить бизнес-анализом</i>	<i>ПК-2.3 Собирает информацию, анализирует, оценивает эффективность проводимого бизнес-анализа в организации</i>	ПК-2.3. З-1. Знает источники информации и аналитические методы подготовки обоснования выбора управленческих решений, связанных с развитием организации ПК-2.3. У-1. Умеет применять информационные технологии для целей управления бизнес-анализом деятельности компании, принимать решения по аналитическому обеспечению разработки стратегии, определению направлений развития компании	Тема 5. Анализ социальных сетей Тема 6. Анализ текстовых данных Тема 7. Нейросетевое моделирование

МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ХАРАКТЕРИЗУЮЩИЕ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ

Перечень учебных заданий на аудиторных занятиях

Полный перечень вопросов для проведения устного опроса студентов по дисциплине «Программные средства бизнес-аналитики»

Индикаторы достижения: ПК-1.2, ПК-2.3

Тема 1: Первичная обработка данных. Графические средства анализа

1. Современные программные пакеты средств статистического анализа.
2. Средства описательной статистики в R, Python: возможности и ограничения.
3. Анализ вариационного ряда в R, Python.
4. Инструменты описательной статистики в R, Python: дисперсия, мода, медиана, стандартная ошибка.
5. Анализ формы распределения в R, Python.
6. Уровень надежности в R, Python. Интерпретация полученных результатов.

Тема 2: Методы классификации. «Случайный лес»

1. Какие существуют подходы при обработке грубых ошибок?
2. Назовите основные методы устойчивого оценивания параметров выборочной совокупности.
3. Приведите пример, в котором резко выделяющее наблюдение обязано своим появлением длинным хвостам распределения, иной модели для части данных, иной модели для всех данных.
4. Приведите примеры распознавания резко выделяющихся наблюдений, опираясь на специальные знания в конкретной области.

Тема 3: Компонентное моделирование. Факторный анализ

1. Что представляет собой формальная постановка задачи и какие данные требуются для иерархической кластеризации?
2. Назовите необходимые исходные данные и постановку задачи для метода К средних.

Тема 4: Ассоциативное моделирование

1. Что понимают под многомерной статистической гипотезой и какие характерные признаки для нее существуют?
2. Приведите примеры многомерных статистических гипотез.
3. Какого рода ошибки могут допускаться и чем определяется достоверность выводов при проверке многомерных статистических гипотез?
4. Будет ли статистической гипотеза:
 - а) о равенстве двух ковариационных матриц, представляющих связи признаков здоровья людей в двух группах: не имеющих хронической заболеваемости и с заболеваемостью?
 - б) о несущественности различий характеров двух человек? И если решать этот же вопрос относительно макроэкономического положения двух стран?
 - в) о том, что многомерная случайная величина $X = (X_1, X_2, X_3)$, представляющая совокупность предприятий, подчиняется нормальному закону распределения?
5. Какими критериями может быть проверена гипотеза о статистическом соответствии двух многомерных выборок?

Тема 5: Анализ социальных сетей

1. Какие особенности должны быть учтены, если для проведения анализа выбирается метод главных компонент или метод главных факторов?
2. Чем отличаются методы ФА от метода главных компонент?
3. Чем объясняется возможность факторного отображения множества элементарных признаков?
4. Как определяется информативность главных компонент? Приведите примеры.
5. Поясните алгоритмическую схему метода Хоттелинга и его преимущества по сравнению с другими методами ФА.

Тема 6: Анализ текстовых данных

1. В чем состоит принципиальное отличие методов многомерных классификаций от комбинационных группировок?
2. Перечислите задачи, решаемые при помощи методов КА.
3. Назовите две основные группы методов КА и укажите их сходство и различие.
4. Какие меры сходства используются при проведении многомерной классификации?
5. Как оценивается качество полученного разбиения многомерной совокупности на кластеры?

Тема 7: Нейросетевое моделирование

1. Приведите известные вам меры связи количественных и не количественных многомерных данных.
2. В каких случаях целесообразно использовать вероятностные коэффициенты связи?

Критерии оценки (в баллах):

Максимальный балл за правильный ответ на устный вопрос преподавателя 0,69 балла.

Неправильный или неполный ответ на вопрос преподавателя 0 баллов

Задания для текущего контроля

ТЕСТЫ

по дисциплине «Программные средства бизнес-аналитики»

Индикаторы достижения: ПК-1.2, ПК-2.3

Тема 1. Первичная обработка данных. Графические средства анализа

- 1) Как переводится DATAMINING?
 - а) “добыча” или “раскопка данных”
 - б) “значение”
 - в) “хранение”
 - г) “перечисление данных”
- 2) Какое требование к переработке информации не верно?
 - а) Данные имеют неограниченный объем
 - б) Данные являются разнородными
 - в) Результаты должны быть конкретны и понятны
 - г) Инструменты для обработки сырых данных должны быть сложны в использовании
- 3) Какая концепция положена в основу современной технологии Data Mining?
 - а) Концепция естествознания

- б) Концепция управления
в) Концепция шаблонов (паттернов)
г) Концепция становления
- 4) Сколько выделяют стандартных типов закономерностей?
а) 4
б) 5
в) 6
г) 7
- 5) Если несколько событий связаны друг с другом, то это...
а) Ассоциация
б) Последовательность
в) Классификация
г) Кластеризация
- 6) Основой для каких систем служит историческая информация, хранящаяся в БД в виде временных рядов?
а) Классификации
б) Последовательности
в) Прогнозирования
г) Ассоциации
- 7) Какой метод используется для оценки параметров уравнения используется, в соответствии с которым $\hat{\beta} = (X^T X)^{-1} X^T Y$, где $\hat{\beta}$ – вектор оценок параметров уравнения, X – матрица независимых переменных, Y – вектор независимых переменных.
а) Метод к-средних;
б) Метод наименьших квадратов;
в) Индексный метод;
г) Иерархические кластер-процедуры.
- 8) По какой формуле определяется наблюдаемое значение F -критерия Фишера (R^2 – множественный коэффициент детерминации, n – число наблюдений, m – число независимых переменных)?
а) $F_{набл} = 1 - \frac{R^2}{n} \frac{n-m}{m}$
б) $F_{набл} = \frac{R^2}{1 - R^2} \frac{n-m}{m}$
в) $F_{набл} = 1 - \frac{R^2}{n} \frac{m-1}{m-1}$
г) $F_{набл} = 1 - \frac{R^2}{n} \frac{m-1}{m}$
- 9) Для оценки надежности регрессионного уравнения наблюдаемое значение F -критерия Фишера должно быть:
а) Больше расчетного значения F -критерия Фишера
б) Меньше расчетного значения F -критерия Фишера
в) Равно расчетному значению F -критерия Фишера
г) F -критерий Фишера не оценивает надежность уравнения.
- 10) Для определения количества значений, находящихся в выборке ниже или выше медианного значения используется:
а) Коэффициент конкордации;
б) Коэффициент корреляции рангов Спирмена;
в) Коэффициент корреляции рангов Кэндалла;
г) Критерий знаков.

11) Какая математическая модель приведена ниже?

$x_{ij} = a_{i1}f_{1j} + a_{i2}f_{2j} + \dots + a_{im}f_{mj} + d_{jv}v_{ij}$, где $i=1, n$; $j=1, m$; $r=1, m$, 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.

- а) Модель факторного анализа;
- б) Математическая модель метода главных компонент;
- в) Мультипликативная модель;
- г) Аддитивная модель.

12) Мультиколлинеарность факторных переменных - это:

- а) Отсутствие связи между факторными переменными;
- б) Тесная связь между факторными переменными;
- в) Многомерная связь между факторными переменными;
- г) Множественная регрессионная модель.

13) Элементами матрицы факторного отображения A по формуле $A = V \Lambda^{-1/2}$ являются:

- а) Линейные коэффициенты корреляции между исходными переменными и главными компонентами
- б) Частные коэффициенты корреляции между исходными переменными и главными компонентами
- в) Множественные коэффициенты корреляции между исходными переменными и главными компонентами

14) Формула расчета расстояния Махаланобиса имеет следующий вид:

б)
$$d(X_i, X_j) = \sqrt{\sum_{k=1}^p (x_i^{(k)} - x_j^{(k)})^2};$$

в)

г)
$$d(X_i, X_j) = \sum_{s=1}^p |x_i^{(s)} - x_j^{(s)}|;$$

$X_i \in S_l, X_j \in S_m$

$$p_{\min}(S_l, S_m) = \min d(X_i, X_j)$$

15) При использовании метода k -средних для классификации многомерных объектов в состав кластера включаются новые объекты таким образом, чтобы внутриклассовая дисперсия:

- а) Стремилась к минимуму;
- б) Стремилась к максимуму;
- в) Оставалась постоянной.

Тема 2. Методы классификации. «Случайный лес»

1) Дискриминантный анализ – совокупность статистических методов многомерной классификации объектов при наличии:

- а) Средних значений;
- б) «Обучающих» выборок;
- в) «Обычных» выборок;
- г) Коэффициентов корреляции.

2) Как интерпретируется DATA MINING?

- а) интеллектуальный анализ данных
- б) значение
- в) хранение
- г) перечисление данных

3) Если несколько событий связаны друг с другом, то это...

- а) Ассоциация
- б) Последовательность

- в) Классификация
- г) Кластеризация
- 4) Data Mining — это процесс обнаружения в сырых данных знаний, необходимых для:
 - а) принятия решений в различных сферах человеческой деятельности
 - б) замены аналитика в процессе принятия решений
 - в) увеличения стоимости анализа данных
 - г) объективных закономерностей
- 5) Если сравнивать Data Mining, машинное обучение и статистику, какая из дисциплин сконцентрирована на едином процессе анализа данных, включает очистку данных, обучение, интеграцию и визуализацию результатов:
 - а) Data Mining
 - б) машинное обучение
 - в) статистика
 - г) Все вышеперечисленное
- 6) “Data Mining” представляет собой направление анализа, основной целью которого является
 - а) поиск нетривиальных закономерностей внутри данных
 - б) разделение данных по определенному признаку
 - в) снижение размерности данных
 - г) исследование взаимосвязей между исходными переменными
- 7) Основой для каких систем не служит историческая информация, хранящаяся в БД в виде временных рядов?
 - а) Классификации
 - б) Последовательности
 - в) Прогнозирования
 - г) Ассоциации
- 8) Data Mining — это процесс обнаружения в сырых данных
 - а) ранее сформулированных гипотез
 - б) неочевидных закономерностей
 - в) практических закономерностей
 - г) большого количества закономерностей
- 9) Выберите характеристику, наиболее подходящую для Data Mining
 - а) подходит для понимания ретроспективных данных
 - б) опирается на ретроспективные данные для получения ответов на вопросы о будущем
 - в) подходит для обобщения ретроспективных данных
 - г) подходит для классификации ретроспективных данных
- 10) Для данных небольшого объема методики Data Mining:
 - а) неприменимы
 - б) применимы
 - в) применимы, но дают неверные результаты
 - г) применимы только для однородных данных
- 11) Закономерности, найденные в процессе использования технологии Data Mining должны обладать такими свойствами:
 - а) быть очевидными
 - б) быть неочевидными
 - в) быть практически полезными
 - г) быть объективными
- 12) Исходные данные для Data Mining:
 - а) представляют собой результат специально организованного статистического наблюдения
 - б) носят спонтанный характер
 - в) засорены большим количеством грубых ошибок и аномальных наблюдений
 - г) как правило, имеют низкую однородность
- 13) Какое требование к переработке информации не верно?
 - а) Данные имеют неограниченный объем

- б) Данные являются разнородными
- в) Результаты должны быть конкретны и понятны
- г) Инструменты для обработки сырых данных должны быть сложны в использовании
- 14) Целесообразность применения методик Data Mining определяется:
 - а) большим объемом исходной информации;
 - б) практической задачей выявления неочевидных закономерностей; наличием
 - в) данных неоднородности;
 - г) все перечисленное
- 15) По логике вычисления средней в совокупности, заведомо засоренной «грубыми ошибками», вес (роль) аномальных наблюдений в формировании среднего значения: а) увеличивается
 - б) уменьшается
 - в) не влияет на среднюю
 - г) аппроксимируется

Тема 3. Компонентное моделирование. Факторный анализ

- 1) В статистике робастным оцениванием называют:
 - а) способы статистического анализа, направленные на снижение влияния на результат аномальных значений и грубых ошибок
 - б) оценку распределений, имеющих отклонения от нормального;
 - в) оценку устойчивости совокупности к внешним и внутренним воздействиям;
 - г) оценку колеблемости данных
- 2) Требованием к проведению дисперсионного анализа является
 - а) равенство групповых дисперсий
 - б) равенство групповых средних
 - в) равные объемы выборки
 - г) все вышеперечисленное
- 3) Для нечисловых данных может применяться
 - а) факторный анализ
 - б) многомерное шкалирование
 - в) компонентное моделирование
 - г) регрессионный анализ
- 4) При проведении цензурирования большого объема данных для исключения подозрительных значений в случае отсутствия нормальности целесообразно применять
 - а) среднее значение
 - б) t-критерий Стьюдента
 - в) правило Томпсона
 - г) подходят все варианты
- 5) При наличии в статистической совокупности аномальных наблюдений и грубых ошибок
 - а) такие наблюдения отбрасываются (исключаются из совокупности и анализ проводится без них)
 - б) такие наблюдения модифицируются (заменяются на значения, близкие к ним);
 - в) такие наблюдения заменяются на среднюю (математическое ожидание);
 - г) все вышеперечисленное;
- 6) Группировка является одним из хороших способов получения робастных оценок. По Вашему мнению, наименьшие потери информации (дисперсии) об исходной совокупности имеет:
 - а) группировка равными интервалами
 - б) группировка равными частотами
 - в) группировка равными вероятностями
 - г) асимптотическая группировка (по F-критерию Фишера)
- 7) Существенные расхождения между сгруппированными и несгруппированными данными могут быть вызваны

- а) засорением выборки аномальными значениями;
 - б) ошибками в расчетах;
 - в) неверным предположением о законе распределения генеральной совокупности;
 - г) все вышеперечисленное
- 8) Дисперсионный анализ при отсутствии нормальности распределения
- а) применим
 - б) неприменим
 - в) применим с ограничениями
 - г) требует преобразования данных
- 9) Наличие нормального распределения требует
- а) факторный анализ
 - б) многомерное шкалирование
 - в) метод главных компонент
 - г) дисперсионный анализ
- 10) Применение методов Data Mining для данных большого объема:
- а) снимает проблему наличия выбросов в данных
 - б) снимает проблему содержательного анализа переменных
 - в) решает проблему мультиколлинеарности в данных
 - г) ничего из перечисленного
- 11) Для поиска нетривиальных закономерностей в данных статистические методы и модели
- а) имеют ограниченное применение (применяются только в исключительных случаях) б) не используются
 - в) применяются наравне со специфическими для Data Mining методами
 - г) не могут быть использованы для данных большого объема
- 12) Методики Data Mining находят широкое применение вследствие:
- а) современных возможностей получения, хранения и обработки большого объема информации
 - б) поиск нетривиальных закономерностей позволяет существенно повысить показатели экономической эффективности
 - в) методики Data Mining являются более совершенными по отношению к традиционным статистическим методам
 - г) только а) и б), но не в)
- 13) Выбросом (аномальным наблюдением) принимается такой объект (элементарная единица совокупности), данные о котором:
- а) являются выбросом хотя бы по одной переменной
 - б) являются выбросом сразу по всем переменным
 - в) содержат пропущенные значения
 - г) возможны а) и б)
- 14) Для выявления структуры и однородных групп в данных, применяется
- а) иерархический кластерный анализ
 - б) метод К средних
 - в) дискриминантный анализ
 - г) метод главных компонент
- 15) Если необходимо максимизировать дисперсию между кластерами, для объединения объектов используется
- а) метод одиночной связи
 - б) метод полных связей
 - в) метод средней связи
 - г) метод Уорда

Тема 4. Ассоциативное моделирование

- 1) Задачей дискриминантного анализа является

- а) выявление существующих внутри совокупности классов, заведомо неизвестных
 - б) определение принадлежности объектов совокупности к той или иной группе
 - в) выявление целостности данных
 - г) снижение размерности данных
- 2) Для кластерного анализа большой объем совокупности
существенен не существенен
должен иметь нормальное распределение не должен содержать выбросов
- 3) Кластерный анализ позволяет проводить
- а) группировку объектов
 - б) группировку признаков
 - в) группировку объектов и группировку признаков
 - г) выделение главных факторов
- 4) Метод Уорда заключается в следующем
- а) на основании матрицы расстояний выбираются ближайшие объекты, которые образуют начальные кластеры, затем в них включаются наиболее близкие объекты, и т. д.
 - б) включение объекта в кластер происходит, если расстояние между объектами не ниже заданного уровня
 - в) каждый кластер изначально состоит из одного объекта; вначале объединяются ближайшие кластеры; дальнейшее объединение кластеров происходит при минимальных приращениях вариации внутри кластеров
 - г) объекты включаются в кластер, если среднее значение близости не ниже заданного порогового уровня
- 5) Функционалом называют
- а) критерий качества разбиения
 - б) формулу расстояния между объектами
 - в) условие оптимальности разбиения
 - г) количество итераций
- 6) В случае, если исследуемые признаки коррелированы, для кластерного анализа применяется
- а) квадрат Евклидова расстояния
 - б) взвешенное Евклидово расстояние
 - в) расстояние «Городских кварталов»;
 - г) расстояние Махаланобиса
- 7) В случае, если необходимо получить наиболее однородные группы, для объединения объектов в кластеры применяется
- а) метод одиночной связи
 - б) метод полных связей
 - в) метод средней связи
 - г) метод Уорда
- 8) Целевая функция в кластерном анализе представляет собой
- а) количество итераций
 - б) критерий качества разбиения
 - в) расстояние между объектами
 - г) условие оптимальности, согласно которому происходит разбиение
- 9) Метод группировки заключается:
- а) в формировании новых классов, если их количество и состав неизвестны;
 - б) в разнесении наблюдений по существующим классам, для которых известны их параметры;
 - в) в ликвидации грубых ошибок и аномальных наблюдений;
 - г) в проверке данных на однородность
- 10) Метод классификации заключается:
- а) в формировании новых классов, если их количество и состав неизвестны;
 - б) в разнесении наблюдений по существующим классам, для которых известны их параметры;

- в) в ликвидации грубых ошибок и аномальных наблюдений;
- г) в проверке данных на однородность
- 11) Задачей иерархического кластерного анализа является
 - а) выявление существующих внутри совокупности классов, заведомо неизвестных
 - б) определение принадлежности объектов совокупности к той или иной группе
 - в) выявление однородности данных
 - г) снижение размерности данных
- 12) В случае, если необходимо получить наиболее компактные группы, для объединения объектов в кластеры применяется
 - а) метод одиночной связи
 - б) метод полных связей
 - в) метод средней связи
 - г) метод Уорда
- 13) Главные факторы содержат информацию
 - а) о самих наблюдениях
 - б) о вариации или дисперсии наблюдений
 - в) об исходных переменных
 - г) о геометрии исходных данных
- 14) Отличие авторегрессионного процесса от процесса скользящей средней состоит в том, что процесс авторегрессии
 - а) имеет лучшую память о предыдущем состоянии
 - б) помнит компонент прошлого случайного шума
 - в) более сглаженный
 - г) все вышеперечисленное
- 15) Какую иерархическую структуру создают деревья решений?
 - а) "ЕСЛИ... ТО..."
 - б) "НИ... НИ..."
 - в) "КОГДА... ТО..."
 - г) "... НИКОГДА..."
- 16) При проверке значимости частного коэффициента корреляции, найденного при $n=16$, используется распределение
 - а) Нормальное
 - б) Стьюдента
 - в) Пирсона
 - г) Пуассона
- 17) В каких пределах изменяется множественный коэффициент корреляции? а) $0 \leq r_{y/xz} \leq 1$
 - б) $-1 \leq r_{y/xz} \leq 1$
 - в) $-\infty \leq r_{y/xz} \leq +\infty$
 - г) $0 \leq r_{y/xz} \leq \infty$
- 18) Задачей логистической регрессии является
 - а) выявление существующих внутри совокупности классов, заведомо неизвестных
 - б) определение принадлежности объектов совокупности к той или иной группе
 - в) выявление целостности данных
 - г) снижение размерности данных
- 19) Главные компоненты содержат информацию
 - а) о самих наблюдениях
 - б) о вариации или дисперсии наблюдений
 - в) об исходных переменных
 - г) о геометрии исходных данных
- 20) Парный коэффициент корреляции между факторами равен 1. Это означает
 - а) наличие нелинейной функциональной связи
 - б) отсутствие связи в) наличие функциональной связи
 - г) отрицательную линейную связь

Тема 5. Анализ социальных сетей

- 1) На основании 20 наблюдений выяснено, что выборочная доля дисперсии случайной величины y , вызванной вариацией x , составит 64%. Чему равен выборочный парный коэффициент корреляции
- а) 0,64
 - б) 0,36
 - в) 0,8
 - г) 0,8 или -0,8
- 2) На основании 20 наблюдений выяснено, что выборочная доля дисперсии случайной величины y , вызванной вариацией x , составит 36%. Чему равен выборочный парный коэффициент корреляции:
- а) 0,36
 - б) 0,16
 - в) 0,6
 - г) + 0,6 или -0,6
- 3) На основании 20 наблюдений выяснено, что выборочная доля дисперсии случайной величины y , вызванной вариацией x , составит 81%. Чему равен выборочный парный коэффициент корреляции:
- а) 0,81
 - б) 0,36
 - в) 0,9
 - г) + 0,9 или -0,9
- 4) Если в данных отчетливо выделяются отдельные группы объектов, то это: а) Ассоциация
- б) Последовательность
 - в) Классификация
 - г) Кластеризация
- 5) Задачей компонентного моделирования является
- а) выявление существующих внутри совокупности классов, заведомо неизвестных
 - б) определение принадлежности объектов совокупности к той или иной группе
 - в) выявление целостности данных
 - г) снижение размерности данных
- 6) Главные факторы, полученные на основе факторного анализа любым некомпонентным методом
- а) линейно неотделимы
 - б) имеют нулевую корреляцию
 - в) коррелированы
 - г) ортогональны
- 7) Имеется таблица сопряженности о количестве предприятий, разделенных по организационной форме и величине (крупные, средние, мелкие). Для анализа взаимосвязей между этими факторами в данном случае не применим
- а) кластерный анализ
 - б) факторный анализ
 - в) дисперсионный анализ
 - г) метод главных компонент
- 8) Для дискриминантного анализа исходные данные
- а) должны быть нормализованы
 - б) должны быть достаточно большого объема в зависимости от количества параметров
 - в) должны быть линейно зависимы
 - г) всё вышеперечисленное
- 9) Применение адаптивных моделей семейств Хольта, Брауна, Уинтерса подразумевает содержательное наличие тренда
- а) верно
 - б) неверно
 - в) необязательно

г) верно, но с поправками при построении моделей

10) Модель авторегрессионного скользящего среднего (ARMA) отличается от модели интегрированного авторегрессионного скользящего среднего (ARIMA), тем, что ARIMA

а) имеет тенденцию оставаться вблизи долгосрочного среднего значения

б) описывает нестационарный процесс

в) обладает улучшенной краткосрочной памятью

г) все вышеперечисленное

11) Многомерное нормальное распределение могут составлять

а) все значения, которые может принимать случайный вектор

б) все значения, которые принимают несколько совместных случайных величин

в) все значения, которые принимает любой набор взаимонезависимых случайных величин

г) все значения, которые принимает любой набор совместных случайных величин

12) У исследователя имеются данные по показателям деятельности нескольких предприятий.

Некоторые из этих показателей являются количественными (величина балансовой прибыли, фондоотдача, показатели устойчивости, ликвидности, и т.д.), некоторые – качественными (организационная форма и пр.). Для выявления сходств отдельных наблюдений и групп предприятий наиболее целесообразно применить

а) дисперсионный анализ

б) факторный анализ

в) шкалирование

г) метод главных компонент

13) Задачей "дерева решений" является

а) выявление существующих внутри совокупности классов, заведомо неизвестных

б) определение принадлежности объектов совокупности к той или иной группе

в) выявление целостности данных

г) снижение размерности данных

14) Задачей "случайного леса" является

а) выявление существующих внутри совокупности классов, заведомо неизвестных

б) определение принадлежности объектов совокупности к той или иной группе

в) выявление целостности данных

г) снижение размерности данных

15) Задачей нейросетевого моделирования является

а) выявление существующих внутри совокупности классов, заведомо неизвестных

б) определение принадлежности объектов совокупности к той или иной группе

в) выявление целостности данных

г) снижение размерности данных

16) Для выделения групп наиболее похожих предприятий по организационной форме и величине лучше всего подходит

а) дисперсионный анализ

б) факторный анализ

в) шкалирование

г) метод главных компонент

17) Для нейросетевого моделирования проблема выбора переменных

а) не существует

б) основана на содержательном анализе

в) решается на основе подгонки модели

г) зависит от кластеризации данных

18) Можно ли интерпретировать нейросетевую модель с точки зрения факторов, влияющих на результативные параметры?

а) можно, если количество скрытых факторов совпадает с числом параметров, влияющих на совокупность

б) нельзя

в) зависит от ситуации

г) скрытые факторы интерпретируются так же, как главные компоненты

- 19) Метод классификации "случайный лес" склонен к "переобучению" в случае:
- а) большого количества исходных переменных
 - б) малого количества исходных наблюдений
 - в) малого количества исходных переменных
 - г) большого количества исходных наблюдений
- 20) Выберите верное утверждение
- а) Нейросетевое моделирование является аналогом работы человеческого мозга;
 - б) Нейросетевые модели позволяют не проводить содержательный анализ исходных факторных признаков;
 - в) Нейросетевые модели не имеют содержательную интерпретацию;
 - г) Нейросетевые модели не могут быть использованы для подгонки "белого шума"

Тема 6. Анализ текстовых данных

- 1) Наиболее употребительными расстояниями и мерами близости между классами объектов являются: (выберите необходимый вариант)
- а) расстояние, измеряемое по принципу "ближайшего соседа"
 - б) расстояние, измеряемого по принципу "дальнего соседа"
 - в) расстояние, измеряемое по принципу "родственной связи"
 - г) расстояние, измеряемое по принципу «значимой связи»
- 2) Оценка значимости моделей семейств Бокса-Дженкинса основывается
- а) на количестве переменных авторегрессии и скользящей средней
 - б) на наличии автокорреляции в регрессионных остатках
 - в) на значимости переменных авторегрессии и скользящей средней
 - г) на подогнанности под поведение данных
- 3) Чему равен коэффициент корреляции показателя самого на себя
- а) 0
 - б) 1
 - в) -1
 - г) ∞
- 4) Многомерное шкалирование решает задачу сходств или различий объектов и представления их в меньшей размерности, основываясь на
- а) максимизации дисперсии вокруг главных осей
 - б) сохранении геометрической структуры данных
 - в) извлечении латентных переменных
 - г) извлечении главных координатных осей
- 5) Коэффициент детерминации между x и y характеризует
- а) - долю дисперсии y , обусловленную влиянием не входящих в модель факторов
 - б) долю дисперсии y , обусловленную влиянием x
 - в) долю дисперсии x , обусловленную влиянием не входящих в модель факторов
 - г) направление зависимости между x и y
- 6) Для какого из перечисленных видов анализа данные можно не нормировать?
- а) кластерный анализ
 - б) дискриминантный анализ
 - в) метод главных компонент
 - г) регрессионный анализ
- 7) Параметры авторегрессии в моделях ARMA/ARIMA содержат информацию:
- а) о предыдущих уровнях ряда динамики
 - б) о предыдущих изменениях ряда динамики
 - в) о предыдущей дисперсии ряда динамики
 - г) о предыдущих значениях "белого шума" ряда динамики
- 8) Параметры скользящей средней в моделях ARMA/ARIMA содержат информацию:
- а) о предыдущих уровнях ряда динамики
 - б) о предыдущих изменениях ряда динамики

- в) о предыдущей дисперсии ряда динамики
- г) о предыдущих значениях "белого шума" ряда динамики
- 9) При декомпозиции ряда динамики с постоянной амплитудой колебаний целесообразно использовать:
 - а) мультипликативную модель
 - б) аддитивную модель
 - в) нелинейную модель
 - г) регрессионную модель
- 10) При декомпозиции ряда динамики с нарастающей/затухающей амплитудой колебаний целесообразно использовать:
 - а) мультипликативную модель
 - б) аддитивную модель
 - в) нелинейную модель
 - г) регрессионную модель
- 11) Модели семейства ARIMA подходят для моделирования:
 - а) стационарных рядов динамики
 - б) нестационарных рядов динамики
 - в) рядов распределения
 - г) ошибок прогнозирования
- 12) Большинство программ, реализующих алгоритм иерархической классификации, предусматривает графическое представление результатов классификации в виде
 - а) дендрограммы
 - б) блок-схемы
 - в) главных компонент
 - г) многомерного распределения
- 13) При применении метода главных компонент
 - а) дисперсия исходного набора максимизируется вдоль главных осей
 - б) дисперсия минимизируется вдоль главных осей
 - в) дисперсия минимизируется вдоль регрессии
 - г) моделируется матрица корреляций
- 14) Случайный вектор можно рассматривать как
 - а) зафиксированные значения нескольких независимых параметров объекта
 - б) многомерную случайную величину
 - в) значения нескольких несовместных случайных величин
 - г) значения нескольких совместных случайных величин
- 15) Опорой случайного вектора называются
 - а) все допустимые исходы наблюдения
 - б) счетное количество исходов наблюдения
 - в) все ненулевые исходы наблюдения
 - г) несчетное количество исходов наблюдения

Тема 7. Нейросетевое моделирование

- 1) Многомерная случайная величина представляет собой
 - а) набор параметров исследуемого объекта, зафиксированных в определенный момент времени
 - б) точка в многомерном пространстве координат
 - в) набор совместных исходов наблюдения одного объекта
 - г) всё вышеперечисленное
- 2) Метод канонических корреляций целесообразно применять в случае
 - а) исходный набор признаков можно разделить на несколько линейных комбинаций
 - б) в исходном наборе признаков встречаются коллинеарные переменные
 - в) число исходных переменных достаточно большое
 - г) все перечисленное

- 3) Канонические переменные
 - а) взаимно не коррелированы
 - б) являются линейными комбинациями исходного набора факторов
 - в) выбираются таким образом, чтобы канонические корреляции были максимальны
 - г) все перечисленное
- 4) Основное назначение метода главных компонент заключается в следующем
 - а) выявление факторов, в наибольшей степени влияющих на распределение
 - б) снижение размерности данных без потери информативности
 - в) построение многофакторной модели совокупности
 - г) всё вышеперечисленное
- 5) Кластерный анализ методом К-средних может применяться:
 - а) для группировки данных
 - б) для классификации данных
 - в) для задач группировки и классификации
 - г) для моделирования данных
- 6) В отличие от логистической регрессии, дискриминантный анализ
 - а) может использоваться для небольшого объема исходной информации
 - б) требует большего объема исходной информации
 - в) применим только для классификации дихотомической переменной
 - г) представляет собой нелинейную модель классификации
- 7) Для оценки значимости канонических корреляций наиболее целесообразно с точки зрения робастности использовать критерий:
 - а) след Пиллаи
 - б) критерий Хи-квадрат Пирсона
 - в) F-критерий Фишера
 - г) лямбда Уилкса
- 8) Какое значение могут принимать канонические корреляции?
 - а) 0,68
 - б) -0,68;
 - в) 1,2
 - г) -1,2
- 9) Факторные нагрузки в каноническом моделировании
 - а) показывают процент объясненной исходной вариации
 - б) показывают процент вариации одного подвектора, который может быть объяснен при наличии второго;
 - в) показывают степень взаимосвязи с исходными переменными;
 - г) показывают процент излишней объясненной вариации
- 10) Критерий чрезмерности в каноническом моделировании
 - а) показывает процент объясненной исходной вариации
 - б) показывает процент вариации одного подвектора, который может быть объяснен при наличии второго;
 - в) показывает степень взаимосвязи с исходными переменными;
 - г) показывает процент излишней объясненной вариации

Критерии оценки (в баллах):

- если тест выполнен в полном объеме (85 и более%)– 4 балла;
- если тест выполнен по большей части (70-84%) – 3 балла;
- если тест выполнен наполовину (50-69%) – 2 балл;
- если тест выполнен в объеме меньше половины (до 50%)– 1 балла;
- если тест выполнен не выполнен вообще – 0 баллов;

Задания для лабораторных работ

Тема 4. Ассоциативное моделирование

Имеются данные о выживших пассажирах Titanic.

По исходным данным, пользуясь методом ассоциативного моделирования, ответьте на следующие вопросы:

Задание 1. Пользуясь знаниями из предыдущей лабораторной работы, выведите общие сведения о данных. Укажите, какие имеются параметры, какого они вида, и какие возможные значения принимают.

Сделайте несколько (4-5) предположений о том, какие закономерности можно найти в этих данных. Для каждого предположения сформулируйте, как должен быть поставлен исследовательский (аналитический) вопрос о его выяснении.

Задание 2.

Выведите правила ассоциации, характеризующие пассажиров 2 класса из левого списка (antecedents).

Задание 3.

Выведите правила, характеризующие выживших, длиной более 2, и отсортируйте их по убыванию критерия интересности (lift).

Сколько всего закономерностей вы получили?

Приведите несколько ассоциативных правил (5-7), представляющих исследовательский интерес.

Интерпретируйте эти правила на основе их содержательного значения либо здравого смысла, с учетом достоверности.

(Например, правило "(Age_Child, Class_2nd) ==> (Survived_Yes)" имеет уровень достоверности 1.0 и подъем 3.0, и интерпретируется как «все дети из класса 2 выжили».)

Приведите несколько ассоциативных правил (3-4), которые не имеют содержательного смысла.

Задание 4.

Мы заметили, что правил по детям из 1 и 3 классов нет. Как следствие, можно неверно интерпретировать правило из предыдущего пункта как то, что дети 2 класса имели больший шанс на выживание, однако, правило всего лишь утверждает, что все дети 2 класса выжили, но не содержит информации о детях других классов, поэтому необходимо получить более подробную информацию. Кроме того, нас не будут интересовать правила длиной менее 2.

Ответьте на вопросы по заданию 4:

А) отличается ли выживаемость детей из 1 класса от ехавших 2-м?

Б) какой процент детей, ехавших первым классом, выжили?

В) какой процент детей, ехавших третьим классом, выжили?

Г) какой процент взрослых, ехавших вторым классом, выжили? Сопоставьте это значение с числом в вопросе (В)

Д) какой процент взрослых, ехавших первым классом, выжили? Сопоставьте это значение с числами в вопросах (В) и (Г)

Задание 5.

Представьте правила в виде графиков. Для этого перестроим список правил

Какие выводы можно сделать из данной диаграммы?

Например, в частности, какой уровень опоры имеют правила с высоким критерием подъема?

Тема 5. Анализ социальных сетей

Имеются данные о сети трех учебных групп по предмету, в которых проявляется проблема академической нечестности (списывание заданий).

(фрагмент исходных данных)

Source,Target

30,84

30,8

30,5

30,26

30,116

30,28

30,31

30,107

45,8

45,94

По результатам анализа напишите аналитическую записку, прояснив, в частности:

1. Понимание ситуации:

- Какие узлы наиболее связаны с другими? Какого рода связь это может быть?
- Кто находится на периферии сети и кто в центре?
- Кто появляется в одних частях сети, но не появляется в других? Почему так может происходить?

2. Анализ деятельности:

- Просматривается ли в сети иерархия? (напр., лидеры)
- Видны ли способы фрагментировать сеть? (например, сосредоточившись на "брокерах")
- Являются ли некоторые модули (подгруппы) более плотными или более разрозненными (то есть потенциально более или потенциально менее подверженными воздействию извне, напр., конкурирующей группой и т.д.)

3. Определение заинтересованных узлов: (либо выявление уязвимостей)

- кто может быть более заинтересованным в информационном наполнении? (более внешние узлы, которые связаны с членами модулей)
- кто из них вероятнее всего задумывается о том, чтобы присоединиться к подпискам? (в анализе реальных и отрицательных, например, преступных сетей здесь целесообразно задать вопрос - могут ли они также привлечь других, но в нашем примере содержательный смысл такого вопроса не столь высокий)

4. Таргетирование

- Какую роль играют отдельные узлы в формировании сети?
- Какие узлы имеют наибольшее число связей с другими?
- Какие узлы уникально соединены с большим количеством других?
- Кто является "брокерами"?
- Какое наиболее ожидаемое влияние на сеть произведет воздействие на отдельные узлы?

Проверьте полученные выводы, письменно ответив на следующие вопросы (сообразно возможностям):

- 1) Совпадают ли результаты с тем, что мы уже знаем? Что в них является необычным?
- 2) Могут ли быть необычные результаты объясняться проблемой с данными?
- 3) Кто может знать ситуацию непосредственно? Каково их мнение?

Критерии оценки (в баллах):

- если творческое задание выполнено в полном объеме (85 и более%)– 10 баллов;
- если творческое задание выполнено по большей части (70-84%) – 7 баллов;
- если творческое задание выполнено наполовину (50-69%) – 5 баллов;
- если творческое задание выполнено в объеме меньше половины (до 50%)– 3 балла;
- если творческое задание не выполнено вообще – 0 баллов;

Задания для творческого рейтинга

Индикаторы достижения: ПК-1.2, ПК-2.3

Тема 1: Первичная обработка данных. Графические средства анализа

Задание 1. Соберите данные и исследуйте описательную статистику по одному из следующих статистических наблюдений, или самостоятельно выбранной области:

- 1) обследование работников промышленных предприятий;
- 2) перепись оборудования на промышленных предприятиях;
- 3) обследование семей рабочих и служащих;
- 4) обследование строительных организаций;
- 5) обследование торговых предприятий;
- 6) изучение спроса на некоторые товары;
- 7) изучение общественного мнения по отдельным вопросам.

По указанному наблюдению определите цель и задачи наблюдения; объекты и единицу наблюдения; основные признаки, подлежащие регистрации; вид, форму и способ наблюдения.

Задание 2. Соберите данные по нескольким (2-4) показателям в интересующей вас сфере, или выбранной из следующего перечня:

- 1) обследование работников промышленных предприятий;
- 2) перепись оборудования на промышленных предприятиях;
- 3) обследование семей рабочих и служащих;
- 4) обследование строительных организаций;
- 5) обследование торговых предприятий;
- 6) изучение спроса на некоторые товары;
- 7) изучение общественного мнения по отдельным вопросам.

По указанному наблюдению определите цель и задачи наблюдения; объекты и единицу наблюдения; основные признаки, подлежащие регистрации; вид, форму и способ наблюдения. Для собранных данных укажите тип данных (временной ряд либо данные об одном временном срезе), по каждой переменной: вид распределения, тип переменных (количественная, качественная, дискретная, непрерывная, номинальная и т.д.)

Пользуясь средствами R, рассчитайте описательные статистики, интерпретируйте показатели и сделайте выводы.

Задание 3. На основе тщательного изучения выборки объемом 868, извлеченной из 11013 хранящихся на складе контейнеров, обнаружено, что 3,6% контейнеров не готовы к отгрузке.

а) определите стандартную ошибку, связанную с приведенной выше оценкой процента, и поясните полученный результат;

б) была бы неожиданной информация, что в действительности 4% из 11013 контейнеров не готовы к отгрузке? Почему? Если 10%? Почему?

Опрос клиентов показал, что из случайно отобранных 200 клиентов 42 недовольны постпродажной поддержкой и обслуживанием.

- а) Вычислите основные статистики: размер выборки n , выборочный процент w , стандартную ошибку доли S_p .
- б) Постройте двусторонний 95% доверительный интервал для процента недовольных среди всех клиентов;
- в) Пусть генеральная совокупность состоит из 28209 клиентов. Рассчитайте количество людей, недовольных постпродажным сервисом, для граничных точек доверительного интервала п. б).
- г) сформулируйте и обоснуйте результаты расчетов.

Задание 4. Часть сборочной линии необходимо регулировать, если консистенция подаваемого пластика становится слишком вязкой или недостаточно вязкой по сравнению со значением вязкости 56,0. Вы решаете производить настройку только тогда, когда есть уверенность, что система “вышла из-под контроля”. Среднее значение вязкости в последних 13 измерениях составило 51,22 со стандартным отклонением 3,18.

1. Сформулируйте нулевую и исследовательскую (альтернативную) гипотезы для двусторонней проверки.
2. Выполните двусторонний тест при уровне значимости 5% и опишите полученный результат.
3. Выполните двусторонний тест при уровне значимости 1% и опишите полученный результат.
4. Сформулируйте, если это возможно, односторонний вывод на основе результата двустороннего теста на уровне значимости 5%.

Задание 5. Пекарня изготавливает хлеб, на этикетках которого указан вес 500 грамм. Ниже приведены значения веса случайно отобранного хлеба (грамм): 510, 480, 490, 505, 500, 510, 515, 495, 480, 520.

1. Определите 95% доверительный интервал для среднего значения веса хлеба 2. Сформулируйте гипотезы H_0 и H_1 .
3. Выполните проверку гипотезы (двусторонняя, уровень 0,05) и поясните полученные результаты.

Тема 2. Методы классификации. «Случайный лес»

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.

Критерии оценки (в баллах):

- если творческое задание по теме выполнено в полном объеме (85 и более%)– 10 баллов;

- если творческое задание по теме выполнено по теме по большей части (70-84%%) – 7 баллов;

- если творческое задание по теме выполнено наполовину (50-69%%) – 5 баллов;

- если творческое задание выполнено в объеме меньше половины (до 50%)– 3 балла;

- если творческое задание не выполнено вообще – 0 баллов;

МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ХАРАКТЕРИЗУЮЩИЕ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ ВО ВРЕМЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Структура экзаменационного билета

<i>Наименование</i>	<i>Максимальное количество баллов</i>
<i>Вопрос 1.</i>	5
<i>Вопрос 2.</i>	8
<i>Практическое задание.</i>	27

Задания, включаемые в экзаменационный билет

Полный 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).
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.

Типовые расчетно-аналитические задания/задачи:

Задача № 1

Рассчитайте сумму регулирующих доходов бюджета области. Расходы бюджета-300 млн. руб. Дефицит бюджета-40 млн. руб. Сумма закрепленных доходов-100 млн. руб.

Задача № 2

Рассчитайте сумму дефицита бюджета области и сумму субвенции. Регулирующие доходы-300 млн. руб. Закрепленные доходы-200 млн. руб. Расходы бюджета-600 млн. руб. Сумма субвенции составляет 30% дефицита бюджета.

Задача № 3

Рассчитайте курс акции и ее рыночную стоимость, если номинальная стоимость акции-1 000 рублей, дивиденд-18%, ссудный процент-14%.

Задача № 4

Рассчитайте размер дивиденда по обыкновенным акциям. Прибыль АО, направляемая на выплату дивидендов, составляет 800 млн. руб. Общая сумма акций-4 500 млн. руб., в том числе привилегированных-1000 млн. руб. с фиксированным размером дивиденда 20% к их номинальной цене.

Задача № 5

Резервы банка равны 3 000 ден. ед. Общая сумма текущих вкладов- 30000 ден. ед. Норма обязательных резервов-20%. Чему равны фактические резервы банка.

Задача № 6

Определите количество денег, необходимое в качестве средства обращения. Сумма цен реализованных товаров-4 500 млрд. руб. Продано с рассрочкой платежа товаров на 42 млрд. руб. Суммы платежей по обязательствам, сроки которых наступили-172 млрд. руб. Сумма взаимопогашающихся платежей-400 млрд. руб. Среднее число оборотов денег за год-10.

Задача № 7

Банк 5 февраля заключил договор срочного банковского вклада на 91 день. Сумма вклада составляет 100 тыс.руб. Процентная ставка 15% годовых, начисленные проценты капитализируются ежедневно, увеличивая сумму вклада. Определите сумму процентного дохода по формуле сложных процентов и сумму вклада с учетом начисленных процентов.

Задача № 8

Используя данные сайта Росбизнесконсалтинг, выявите тенденцию процентных ставок за последние 12 месяцев, рассчитайте среднюю процентную ставку и степень колеблемости процентных ставок по межбанковским кредитам.

Задача № 9

Используя данные сайта Росбизнесконсалтинг, выявите тенденцию процентных ставок за последние 12 месяцев, рассчитайте среднюю процентную ставку и степень колеблемости процентных ставок по срочным депозитам.

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

Шкала оценивания		Формируемые компетенции	Индикатор достижения компетенции	Критерии оценивания	Уровень освоения компетенций
85 – 100 баллов	«отлично»	ПК-1 Способен обосновывать подходы, используемые в бизнес-анализе	ПК-1.2 Определяет подход к работе с информацией бизнес-анализа	Знает верно и в полном объеме: источники информации и аналитические методы подготовки обоснования выбора управленческих решений; Умеет верно и в полном объеме: работать с базами данных, интернет-ресурсами, распространенными программами для бизнеса, подготовки сводных, аналитических отчетов для управления компанией с учетом поставленной цели и задач;	Продвинутый
		ПК-2 Способен руководить бизнес-анализом	ПК-2.3 Собирает информацию, анализирует, оценивает эффективность проводимого бизнес-анализа в организации	Знает верно и в полном объеме: источники информации и аналитические методы подготовки обоснования выбора управленческих решений, связанных с развитием организации Умеет верно и в полном объеме: применять информационные технологии для целей управления бизнес-анализом деятельности компании, принимать решения по аналитическому обеспечению разработки стратегии, определению направлений развития компании	
70 – 84 баллов	«хорошо»	ПК-1 Способен обосновывать подходы, используемые в бизнес-анализе	ПК-1.2 Определяет подход к работе с информацией бизнес-анализа	Знает с незначительными замечаниями: источники информации и аналитические методы подготовки обоснования выбора управленческих решений; Умеет с незначительными замечаниями: работать с базами данных, интернет-ресурсами, распространенными программами для бизнеса, подготовки сводных, аналитических отчетов для управления компанией с учетом поставленной цели и задач;	Повышенный
		ПК-2 Способен	ПК-2.3 Собирает	Знает с незначительными замечаниями: источники информации	

		руководить бизнес-анализом	информацию, анализирует, оценивает эффективность проводимого бизнес-анализа в организации	и аналитические методы подготовки обоснования выбора управленческих решений, связанных с развитием организации Умеет с незначительными замечаниями: применять информационные технологии для целей управления бизнес-анализом деятельности компании, принимать решения по аналитическому обеспечению разработки стратегии, определению направлений развития компании	
50 – 69 баллов	«удовлетворительно»	ПК-1 Способен обосновывать подходы, используемые в бизнес-анализе	ПК-1.2 Определяет подход к работе с информацией бизнес-анализа	Знает на базовом уровне, с ошибками: источники информации и аналитические методы подготовки обоснования выбора управленческих решений Умеет на базовом уровне, с ошибками: работать с базами данных, интернет-ресурсами, распространенными программами для бизнеса, подготовки сводных, аналитических отчетов для управления компанией с учетом поставленной цели и задач;	Базовый
		ПК-2 Способен руководить бизнес-анализом	ПК-2.3 Собирает информацию, анализирует, оценивает эффективность проводимого бизнес-анализа в организации	Знает на базовом уровне, с ошибками: источники информации и аналитические методы подготовки обоснования выбора управленческих решений, связанных с развитием организации Умеет на базовом уровне, с ошибками: применять информационные технологии для целей управления бизнес-анализом деятельности компании, принимать решения по аналитическому обеспечению разработки стратегии, определению направлений развития компании	
менее 50 баллов	«неудовлетворительно»	ПК-1 Способен обосновывать подходы, используемые в бизнес-анализе	ПК-1.2 Определяет подход к работе с информацией бизнес-анализа	Не знает на базовом уровне: источники информации и аналитические методы подготовки обоснования выбора управленческих решений; Не умеет на базовом уровне: работать с базами данных, интернет-ресурсами, распространенными программами для бизнеса, подготовки сводных, аналитических отчетов для управления компанией с учетом поставленной цели и задач;	Компетенции не сформированы
		ПК-2 Способен руководить	ПК-2.3 Собирает информацию, анализирует,	Не знает на базовом уровне: источники информации и аналитические методы подготовки обоснования выбора управленческих	

		бизнес-анализом	оценивает эффективность проводимого бизнес-анализа в организации	решений, связанных с развитием организации Не умеет на базовом уровне: применять информационные технологии для целей управления бизнес-анализом деятельности компании, принимать решения по аналитическому обеспечению разработки стратегии, определению направлений развития компании	
--	--	-----------------	--	--	--