

**Федеральное государственное образовательное бюджетное учреждение
высшего образования «ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ
ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)**

**Владикавказский филиал Финуниверситета
Кафедра «Математика и информатика»**

Утверждаю
Директор Владикавказского
филиала Финуниверситета

Т.А. Хубаев
« 30 » июня 2023 г.

Волик М.В.

Управление данными предприятия

Оценочные материалы для проведения текущего контроля успеваемости и
промежуточной аттестации студентов, обучающихся по
направлению подготовки: 38.03.05 Бизнес-информатика,
образовательная программа «Цифровая трансформация управления
бизнесом», профиль «ИТ-менеджмент в бизнесе»

*Одобрено заседанием кафедры «Математика и информатика»
(протокол № 12 от 25 мая 2023 г.)*

Владикавказ 2023

Содержание

1. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине.....	3
2. Показатели и критерии оценивания компетенций на различных этапах их формирования, описание шкал оценивания	3
3. Оценочный лист результатов обучения к дисциплине	6
4. Контрольные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний	7

1. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с компетенциями/индикаторами достижения компетенции
ПКН-1	Способность внедрять транзакционные системы и консультировать по вопросам систем сбора, накопления и хранения транзакционных данных	1. Проводит анализ рынка систем сбора, накопления и хранения транзакционных данных.	Знать: современное состояние рынка систем сбора, накопления и хранения транзакционных данных и перспектив их использования в организации. Уметь: использовать современные технологии для решения поставленных.
		2. Внедряет системы сбора, накопления и хранения транзакционных данных.	Знать: особенности внедрения системы сбора, накопления и хранения транзакционных данных. Уметь: использовать современные технологии для решения поставленных.
		3. Консультирует по вопросам применения системы сбора, накопления и хранения транзакционных данных.	Знать: современное состояние основных транзакционных систем. Уметь: использовать современные технологии для решения поставленных.

2. Показатели и критерии оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Уровень сформированности компетенций			
«Минимальный уровень не достигнут»	«Минимальный уровень»	«Средний уровень»	«Высокий уровень»
Компетенции не сформированы.	Компетенции сформированы.	Компетенции сформированы.	Компетенции сформированы.
– Знания отсутствуют, умения и навыки не	– Сформированы базовые структуры знаний.	– Знания обширные, системные. – Умения носят	– Знания твердые, аргументированные, всесторонние.

сформированы.	– Умения фрагментарны и носят репродуктивный характер. – Демонстрируется низкий уровень самостоятельности практического навыка.	репродуктивный характер, применяются к решению типовых заданий. – Демонстрируется достаточный уровень самостоятельности устойчивого практического навыка.	– Умения успешно применяются к решению как типовых, так и нестандартных творческих заданий. – Демонстрируется высокий уровень самостоятельности, высокая адаптивность практического навыка
Описание критериев оценивания			
Обучающийся демонстрирует: – существенные пробелы в знаниях учебного материала; – допускаются принципиальные ошибки при ответе на основные вопросы билета, отсутствует знание и понимание основных понятий и категорий; – непонимание сущности дополнительных вопросов в рамках заданий билета; – отсутствие умения выполнять практические задания, предусмотренные программой дисциплины; – отсутствие готовности (способности) к дискуссии и низкую степень контактности.	Обучающийся демонстрирует: – знания теоретического материала; – неполные ответы на основные вопросы, ошибки в ответе, недостаточное понимание сущности излагаемых вопросов; – неуверенные и неточные ответы на дополнительные вопросы; – недостаточное владение литературой, рекомендованной программой дисциплины; – умение без грубых ошибок решать практические задания, которые следует выполнить.	Обучающийся демонстрирует: – знание и понимание основных вопросов контролируемого объема программного материала; – твердые знания теоретического материала. – способность устанавливать и объяснять связь практики и теории, выявлять противоречия, проблемы и тенденции развития; – правильные и конкретные, без грубых ошибок, ответы на поставленные вопросы; – умение решать практические задания, которые следует выполнить; – владение основной литературой, рекомендованной программой	Обучающийся демонстрирует: – глубокие, всесторонние и аргументированные знания программного материала; – полное понимание сущности и взаимосвязи рассматриваемых процессов и явлений, точное знание основных понятий в рамках обсуждаемых заданий; – способность устанавливать и объяснять связь практики и теории; – логически последовательные, содержательные, конкретные и исчерпывающие ответы на все задания билета, а также дополнительные вопросы экзаменатора; – умение решать практические

		дисциплины; – наличие собственной обоснованной позиции по обсуждаемым вопросам. – возможны незначительные оговорки и неточности в раскрытии отдельных положений вопросов билета, присутствует неуверенность в ответах на вопросы	задания; – свободное использование в ответах на вопросы материалов рекомендованной основной и дополнительной литературы.
Оценка «неудовлетворительно» / «не зачтено»	Оценка «удовлетворительно» / «зачтено»	Оценка «хорошо» / «зачтено»	Оценка «отлично» / «зачтено»

3. Оценочный лист результатов обучения к дисциплине

Примерные виды работ обучающегося, формирующие текущий контроль по дисциплине

№ п/п	Вид учебной деятельности	Баллы	Максимум за семестр (модуль)
1.	Ведение конспекта лекции	0,15	3
2.	Активное вовлечение в интерактивный процесс	0,1	5
3.	Реферативный / аналитический обзор дополнительного материала по теме	0,5	5
4.	Презентация по заданной теме (разделу)	0,3	3
5.	Выполнение практических и самостоятельных работ	0,5	19
6.	Выполнение вида текущего контроля	5	5
	Всего за семестр (модуль)		40

Критерии балльной оценки включают качество подготовки обучающихся к семинарским занятиям, выполнения различных видов самостоятельной работы (при этом результаты регулярно проводимого экспресс-контроля успеваемости обучающихся должны составлять не менее 50% балльной оценки), а также посещение аудиторных занятий (не более 15% балльной оценки).

При подведении итогов текущего контроля успеваемости в середине семестра его результаты в 20-балльной оценке вносятся в ведомости текущего контроля успеваемости и использованием ТЭКУ (в филиалах – в ведомости текущего контроля успеваемости) и на соответствующей странице предметно-группового журнала преподавателям, ведущим аудиторные занятия.

Обучающийся, получивший от 7 до 20 баллов, считается аттестованным, получивший от 0 до 6 баллов – не аттестованным.

Результаты текущего контроля успеваемости за вторую половину семестра / в модуле преподаватели, ведущие аудиторные занятия, сообщают обучающимся на последнем занятии и вносят в ведомости (в колонку «баллы, работа в семестре») в день проведения зачета/экзамена.

Итоговая сумма баллов, полученная обучающимся по результатам текущего контроля успеваемости в семестре/модуле и промежуточной аттестации, преобразуется в пятибалльную систему оценок.

Порядок перевода 100-балльной оценки в пятибалльную

100-балльная система	5-балльная система
86-100	отлично
70-85	хорошо
50-69	удовлетворительно
менее 50	неудовлетворительно
50-100	зачтено
менее 50	не зачтено

**4. Контрольные материалы, необходимые для оценки индикаторов
достижения компетенций, умений и знаний**

Текущий контроль:

Тема 1. Характеристики больших данных (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Основы управления организацией (предприятием).
- Предприятие как объект и субъект управления.
- Функции, уровни и общие принципы организации управления предприятием.
- Распределение полномочий, власть и ответственность.
- Структура управления предприятием.

Задание 1. Проведите анализ текущего состояния рынка систем сбора, накопления и хранения транзакционных данных, рассматривая основные компании на российском рынке. Охарактеризуйте тенденции и направления развития рынка: интеграция с BI-системами и с инструментами машинного обучения, разработка инструментов предиктивной аналитики.

Тема 2. Платформа управления данными Ростелекома (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Создание кластера RT.DataLake.

- Аналитика SQL: Агрегирование, оконные функции.

Задание 1. Проведите обзор функций платформы управления данными Ростелекома.

Тема 3. Стек технологий Arenadata (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Рекурсивные запросы (конструкция WITH).
- Соединение запросов, как множеств (UNION, INTERSECT, EXCEPT).

Задание 1. Проведите анализ текущего состояния мирового рынка систем сбора, накопления и хранения транзакционных данных. Охарактеризуйте тенденции и направления развития рынка: интеграция с BI-системами и с инструментами машинного обучения, разработка инструментов предиктивной аналитики.

Тема 4. Инструменты и технологии сбора, хранения, обработки и анализа больших данных VK Cloud Solutions (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Применение математических XYZ-анализа.

Задание 1. Организуйте облачное развертывание системы сбора, накопления и хранения транзакционных данных, используя современные технические средства стека технологий Arenadata.

Задание 2. Организуйте облачное развертывание системы сбора, накопления и хранения транзакционных данных, используя современные технические средства Yandex Cloud.

Тема 5. Инструменты и технологии сбора, хранения, обработки и анализа больших данных Cloud (Sber) (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Многомерный ABC-XYZ анализ на SQL.

Задание 1. Для выбранного предприятия определите области, в которых могут быть использованы систем сбора, накопления и хранения транзакционных данных. Укажите возможные направления их внедрения и

эффект от их использования.

Тема 6. Инструменты и технологии сбора, хранения, обработки и анализа больших данных Yandex Cloud (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- RFM-анализ на больших объемах структурированных данных.

Задание 1. Для выбранного предприятия проведите консультацию заказчика по вопросам применения систем сбора, накопления и хранения транзакционных данных.

Тема 7. Инструменты и технологии сбора, хранения, обработки и анализа больших данных Selectel (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Платформа для аналитики и работы с Big Data на вычислительных мощностях Selectel.

Задание 1. На основе анализа информационных потоков организации, предложите различные варианты разработки хранилища данных «под ключ» и создания рекомендательных сервисов для оптимизации производственных процессов.

Тема 8. Машинное обучение для анализа больших данных (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Машинное обучение для решения задач Data Mining.

Задание 1. В целях обеспечения импортозамещения технологий обработки и хранения данных, внедрения процесса управления данными в организации, обеспечения горизонтального масштабирования, повышения качества управления, организации эффективной работы внутрикорпоративных структур, предложите различные варианты развертывания платформы управления данными у ведущих облачных провайдеров, необходимых для построения, создания масштабных озер данных и реализации алгоритмов искусственного интеллекта на базе моделей Data Science.

Тема 9. Распределенные вычисления Apache Spark (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Анализ больших данных с использованием фреймворка Spark.

Задание 1. Для выбранного предприятия выявите необходимых изменений в информационных потоках организации, определите мероприятия адаптации ИТ-инфраструктуры для внедрения платформы управления данными. Разработайте предложения для заказчика по выбору направлений изменений ИТ-ландшафта предприятия.

Тема 10. Аналитика данных в SQLite (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Подключение к базе данных SQLite из SQL-клиента (DBeaver).

Задание 1. Для выбранного предприятия определите необходимый состав компонентов платформы управления данными, проведите проектирование целевой архитектуры информационных потоков организации. Разработайте предложения для заказчика по выбору направлений изменений ИТ-ландшафта.

Тема 11. OLAP in-memory database DuckDB (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Аналитика данных в OLAP in-memory database DuckDB.

Задание 1. Организуйте облачное развертывание системы сбора, накопления и хранения транзакционных данных, используя современные технические средства.

Тема 12. Унифицированный интерфейс для распределенных вычислений Fugue (ПКН-1 (1, 2, 3))

Вопросы для обсуждения:

- Использование FugueSQL для определения сквозных рабочих процессов.

Задание 1. Подготовьте описание сквозных рабочих процессов компании по выбору с использованием FugueSQL.

Тема 13. Apache Superset

Вопросы для обсуждения:

- Применение платформы Apache Superset для визуального анализа и исследования данных.

Задание 1. Подготовьте визуальный анализ и исследование данных компании по выбору с использованием Apache Superset.

Примерный перечень заданий расчетно-аналитической работы (ПРН-1 (1, 2, 3))

Студенту в начале семестра предлагается выбрать кейс, на основе которого разрабатываться аналитический проект. Для этого можно использовать датасеты с Kaggle , содержащие клиентские данные (<https://www.kaggle.com/datasets?search=customer> - всего 2466 различных Datasets).

Задание:

1. Импортировать данные в BigQuery. Обогащать датасет, проводя клиентскую аналитику: выполнить ABC-XYZ, RFM, когортный -анализ.
2. Создать, обучить, оценить по метрикам качества, объяснить модели машинного обучения используя SQL; создать интерактивный отчет (дашборд) в Looker Studio.

Примерный перечень заданий контрольной работы (ПРН-1 (1, 2, 3))

На данных кейса предыдущего семестра:

1. В Apache Zeppelin или Colab провести исследовательский анализ данных на Spark, создать и оценить модель машинного обучения.
2. Провести ABC-XYZ, RFM, когортный -анализ в OLAP in-memory database DuckDB.
3. Создать несколько дашбордов в RT.DataVision с использованием интегрированной средой разработки SQL, выявить инсайты, настроить встраивание контента RT.DataVision во внешние сервисы.

Промежуточная аттестация

Примерные вопросы для подготовки к зачету (ПКН-1 (1, 2, 3))

1. Охарактеризуйте понятие «Big Data», источники и характеристики больших данных.
2. Охарактеризуйте принципы работы с большими данными: горизонтальная масштабируемость, отказоустойчивость, локальность данных.
3. Охарактеризуйте стек (экосистему) инструментов Hadoop.
4. Охарактеризуйте программную модель MapReduce, поток данных в MapReduce.
5. Охарактеризуйте проект Apache Spark, преимущества Spark и назначение модулей Spark.
6. Охарактеризуйте облачные технологии обработки больших данных.
7. Охарактеризуйте отличия озера данных и хранилища данных.
8. Охарактеризуйте платформа управления данными Ростелекома.
9. Охарактеризуйте стек технологий Arenadata.
10. Охарактеризуйте инструменты и технологии сбора, хранения, обработки и анализа больших данных VK Cloud Solutions.
11. Охарактеризуйте инструменты и технологии сбора, хранения, обработки и анализа больших данных Cloud (Sber).
12. Охарактеризуйте инструменты и технологии сбора, хранения, обработки и анализа больших данных Yandex Cloud.
13. Охарактеризуйте инструменты и технологии сбора, хранения, обработки и анализа больших данных Selectel.
14. Охарактеризуйте возможности машинного обучения для анализа больших данных в Google BigQuery.

Примерные вопросы для подготовки к экзамену (ПКН-1 (1, 2, 3))

1. Охарактеризуйте источники больших данных: веб-сайты и мобильные приложения, корпоративные приложения, IoT.

2. Охарактеризуйте экосистему Hadoop
3. Охарактеризуйте ключевые возможности HDFS
4. Охарактеризуйте технологию Map Reduce
5. Охарактеризуйте ключевые возможности Hive
6. Охарактеризуйте ключевые возможности Impala
7. Охарактеризуйте ключевые возможности Apache Spark
8. Охарактеризуйте Apache Spark MLlib
9. Охарактеризуйте ключевые возможности Apache Kafka
10. Охарактеризуйте ключевые возможности Apache Presto и HBase
11. Охарактеризуйте ключевые возможности Apache Flink
12. Охарактеризуйте ключевые возможности Apache Beam
13. Охарактеризуйте понятие и назначение Data Lakes
14. Охарактеризуйте ключевые возможности Amazon Web Services S3 (Simple Storage Service)
15. Охарактеризуйте ключевые возможности Google BigQuery
16. Охарактеризуйте ключевые возможности DataLens для аналитики Big Data
17. Охарактеризуйте ключевые возможности Google Data Studio для аналитики Big Data
18. Охарактеризуйте ключевые возможности продукты AutoML (на примере любого облачного провайдера)
19. Охарактеризуйте ключевые возможности аналитики Big Data в публичном облаке VK Cloud Solutions
20. Охарактеризуйте ключевые возможности ClickHouse Database
21. Охарактеризуйте ключевые возможности Greenplum Database
22. Охарактеризуйте возможности сервиса Яндекс.Облака Yandex Data Proc
23. Охарактеризуйте ключевые возможности Yandex Data Science Virtual Machine

24. Охарактеризуйте ключевые возможности MapReduce Service Sbercloud
25. Охарактеризуйте ключевые возможности Data Lake HDInsight Sbercloud
26. Охарактеризуйте ключевые возможности Dataproc Sbercloud
27. Охарактеризуйте ключевые возможности ML Spase Sbercloud
28. Data Mining и основные типы.
29. Охарактеризуйте метрики качества модели классификации.
30. Охарактеризуйте метрики качества задачи кластерного анализа.
31. Охарактеризуйте метрики качества задачи регрессии.
32. Оценка качества обучения модели.

Пример содержания экзаменационного билета

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № _____

1. (10 баллов) Охарактеризуйте ключевые возможности ML Spase Sbercloud. (ПКН-1 (1, 2, 3))
2. (20 баллов) По индивидуальному заданию (кейс №...): импортировать данные в BigQuery. Обогащать датасет, проводя клиентскую аналитику: выполнить ABC-XYZ, RFM, когортный -анализ. (ПКН-1 (1, 2, 3))
3. (30 баллов) Создать несколько дашбордов в RT.DataVision с использованием интегрированной средой разработки SQL, выявить инсайты, настроить встраивание контента RT.DataVision во внешние сервисы. (ПКН-1 (1, 2, 3))