


ФГБОУ ВО НОВОСИБИРСКИЙ ГАУ
Кафедра информационных технологий и моделирования

Рег. № ФМЧУ.03-05013
«5» 10.03.2022г.

УТВЕРЖДЕН
 на заседании кафедры
 Протокол от «23» 03 2022 г. № 2
 Заведующий кафедрой информационных
 технологий и моделирования
 О.В. Арафонова
 (подпись)

ФОНД
ОЦЕНОЧНЫХ СРЕДСТВ

Б1.О.05 Технологии обработки и анализа бизнес-данных
 Цифры и наименования дисциплины
 38.04.01 Экономика
 Код и наименование направления подготовки
 Финансовый менеджмент в услугах цифровизации
 Направленность (профиль)

Паспорт
фонда оценочных средств

№ п/п	Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
1	Вводная лекция	ОПК-3	Вопросы для коллоквиума
2	Python для анализа данных	ОПК-3	Вопросы для коллоквиума
3	Библиотеки для анализа данных: Numpy, Matplotlib, Pandas	ОПК-3	Вопросы для коллоквиума
4	Искусственный интеллект	ОПК-3	Вопросы для коллоквиума
5	Машинное обучение	ОПК-3	Вопросы для коллоквиума
	Контрольная работа, зачет с оценкой	ОПК-3	Задание для контрольной работы, вопросы к зачету

Тема 1: Вводная лекция.

1. Что такое Data Science;
2. Кто такой Data Scientist;
3. Какие прогнозы делает Data Scientist;
4. Итоги работы дата-сайентиста;
5. Польза от прогнозов;
6. Бизнес аналитика и Data Science;
7. Data Scientist в бизнесе, в банках, в спортивных компаниях в IT-сфере, на производстве, в страховых компаниях, в медицине, в сельскохозяйстве.

Критерии оценки:

Для оценки работы вводится 5 балльная оценочная шкала. На заключительном этапе оценочная шкала переводится в «зачтено» или «не зачтено».

Оценочная шкала для итоговой проверки работы заключается в следующем:

1. Для отметки «Зачтено» необходимо набрать свыше 3 баллов;
2. Для отметки «Не зачтено» – количество баллов от 0 до 3.

Шкала распределения баллов для оценки работы

Количество баллов	Оценка в баллах			
	Правильность ответа на основной вопрос	Ответ на каждый дополнительный вопрос	Максимальное количество баллов при ответе на дополнительные вопросы	Итого баллов по вопросу
3		1	2	5

Тема 2: Python для анализа данных

1. Структуры данных;
2. Работа с индексами;
3. Чистка данных;
4. Математически операции;
5. Операции объединения данных;
6. Агрегирование данных;
7. Визуализация данных;
8. Google Colabotory.

Критерии оценки:

Для оценки работы вводится 5 балльная оценочная шкала. На заключительном этапе оценочная шкала переводится в «зачтено» или «не зачтено».

Оценочная шкала для итоговой проверки работы заключается в следующем:

1. Для отметки «Зачтено» необходимо набрать свыше 3 баллов;
2. Для отметки «Не зачтено» – количество баллов от 0 до 3.

Шкала распределения баллов для оценки работы

Количество баллов	Оценка в баллах			
	Правильность ответа на основной вопрос	Ответ на каждый дополнительный вопрос	Максимальное количество баллов при ответе на дополнительные вопросы	Итого баллов по вопросу
3		1	2	5

Тема 3: Библиотеки для анализа данных: Numpy, Matplotlib, Pandas

1. Numpy.
2. Matplotlib.
3. Pandas.
4. Основные объекты для анализа данных.

Критерии оценки:

Для оценки работы вводится 5-балльная оценочная шкала. На заключительном этапе оценочная шкала переводится в «зачтено» или «не зачтено».

Оценочная шкала для итоговой проверки работы записывается в следующем:

1. Для отметки «Зачтено» необходимо набрать свыше 3 баллов.
2. Для отметки «Не зачтено» – количество баллов от 0 до 3.

Шкала распределения баллов для оценки работы

Количество баллов	Оценка в баллах			
	Правильность ответа на основной вопрос	Ответ на каждый дополнительный вопрос	Максимальное количество баллов при ответе на дополнительные вопросы	Итого баллов по вопросу
3		1	2	5

Тема 4: Искусственный интеллект

1. Особенности искусственного интеллекта, основные понятия и задачи области машинного обучения, принципы обучения алгоритмов, примеры применения в бизнесе.
2. История и современного интеллекта.
3. Области искусственного интеллекта.
4. Технологии работы с большими данными.
5. Интеграция в бизнес-процессы.
6. Направления применения. Успехи применения.
7. Методология управления проектами по анализу данных.

Критерии оценки:

Для оценки работы вводится 5-балльная оценочная шкала. На заключительном этапе оценочная шкала переводится в «зачтено» или «не зачтено».

Оценочная шкала для итоговой проверки работы записывается в следующем:

1. Для отметки «Зачтено» необходимо набрать свыше 3 баллов.
2. Для отметки «Не зачтено» – количество баллов от 0 до 3.

Шкала распределения баллов для оценки работы

Количество баллов	Оценка в баллах			
	Правильность ответа на основной вопрос	Ответ на каждый дополнительный вопрос	Максимальное количество баллов при ответе на дополнительные вопросы	Итого баллов по вопросу
3		1	2	5

Тема 5: Машинное обучение

1. Методы машинного обучения.
2. Классификация. Задача классификации.
3. Табличные данные. Создание алгоритма классификации.
4. Регрессия. Линейные модели.
5. Переобучение. Метрики классификации и регрессии.
6. Метрики качества регрессии.
7. Метрики качества классификации.
8. Классификация. Задача классификации.
9. Алгоритмы классификации.
10. Понижение размерности.
11. Выбор признака. Выделение признаков.
12. Визуализация данных.
13. Ассоциации и рекомендательные системы. Ассоциативные правила.
14. Рекомендательные системы.
15. Обучение с подкреплением.
16. Кул-матрица наград.
17. Алгоритмы обучения с подкреплением.
18. Ансамблевые методы: стеклинг, бэггинг, бустинг.
19. Решающие деревья.
20. Ансамблирование. Визуализация.

Критерии оценки:

Для оценки работы вводятся 5 балльная оценочная шкала. На заключительном этапе оценочная шкала переводится в «зачтено» или «не зачтено».

Оценочная шкала для итоговой проверки работы заключается в следующем:

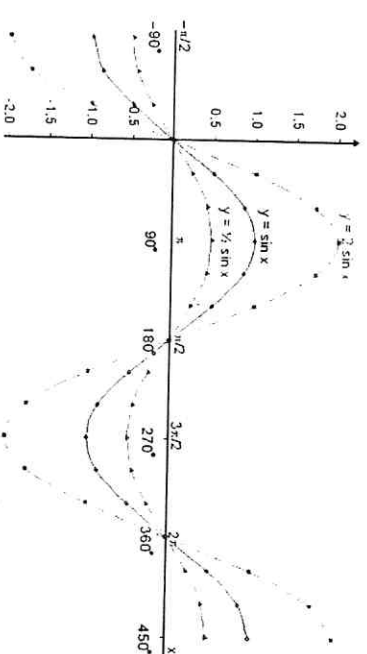
1. Для отметки «Зачтено» необходимо набрать выше 3 баллов.
2. Для отметки «Не зачтено» – количество баллов от 0 до 3.

Шкала распределения баллов для оценки работы

Количество баллов	Оценка в баллах			
	Правильность ответа на основной вопрос	Ответ на каждый дополнительный вопрос	Максимальное количество баллов при ответе на дополнительные вопросы	Итого баллов по вопросу
3		1	2	5

Задание для контрольной работы

1. Нарисуйте как можно более похожую картинку в matplotlib. Добавьте к ней описание осей, сетку, легенду, название. Сохраните картинку из Jupyter или Colab и отправьте преподавателю.



2. Задание выполняется по данным, собранным в ходе своей научно-исследовательской работы.

Каждому из них необходимо:

Задание 1

Прочитать файл в датафрейм, так, чтобы появились корректные колонки без каких-либо дополнительных действий, и отобразите первые 5 строчек.

Задание 2

Рассчитать показатели, позволяющие провести анализ данных. Сформулировать выводы.

Задание 3

С помощью pairplot построить попарные графики с раскраской по любому полю.

Задание 4

Построить любые интересные графики, позволяющие расширить анализ данных. Объяснить почему вы построили именно их.

Задание 5

Постройте сводную таблицу, используя метод pivot table.

Задание 6

Попробуйте найти еще какие-нибудь зависимости в своих данных с помощью доступных вам сейчас методов.

Задание 7

Используйте выборки, агрегацию, графики. Сформулируйте выводы по проделанной работе.

Критерии оценки:

- Ометка «Зачено» выставляется, если задание выполнено.
- Ометка «Не зачено» выставляется, если задание не выполнено.

ФГБОУ ВО Новосибирский ГАУ

Кафедра информационных технологий и моделирования

Вопросы к зачету с оценкой

Полностью прочтите. Бизнес-аналитика и Data Science.

Data Scientist в бизнесе, в банках, в спортивных компаниях в IT-сфере, на производстве, в страховых компаниях, в медицине, в сельском хозяйстве.

1. Структуры данных.
2. Работа с индексами.
3. Чистка данных.
4. Математические операции.
5. Операции объединения данных.
6. Агрегации данных.
7. Визуализация данных.
8. Numpy, Matplotlib, Pandas.
9. Основные объекты для анализа данных.
10. Особенности искусственного интеллекта, основные понятия и задачи области машинного обучения, принципы обучения алгоритмов, примеры применения в бизнесе.
11. История и искусственного интеллекта.
12. Области и искусственного интеллекта.
13. Технологии работы с большими данными.
14. Интеграция в бизнес-процессы.
15. Направления применения. Успехи применения.
16. Методология управления проектами по анализу данных.
17. Методы машинного обучения.
18. Классификация. Задача классификации.
19. Создание алгоритма классификации.
20. Регрессия. Линейные модели. Переобучение.
21. Метрики классификации и регрессии.
22. Метрики качества регрессии.
23. Метрики качества классификации.
24. Классификация. Задача классификации.
25. Алгоритмы классификации.
26. Обзор признаков. Выделение признаков.
27. Визуализация данных.
28. Ассоциации и рекомендательные системы. Ассоциативные правила.
29. Рекомендательные системы.
30. Обучение с подкреплением.
31. Кулинарная наука.
32. Алгоритмы обучения с подкреплением.
33. Ансамблевые методы: стеклинг, бэггинг, бустинг.
34. Решающие деревья. Ансамблирование. Виды ансамблей.

Критерии оценки:

Отметка «ОТЛИЧНО» - дан полный, развернутый ответ на поставленные вопросы, показана совокупность осознанных знаний по дисциплине; в ответе прослеживается четкая структура и логическая последовательность, отражающая сущность раскрываемых понятий. Ответ изложен литературным языком с использованием современной терминологической терминологии. Могут быть допущены незначительные опечатки в написании понятий, и исправленные студентом самостоятельно в процессе ответа.

Отметка «ХОРОШО» - дан полный, развернутый ответ на поставленный вопрос. Ответ четко структурирован, логичен, изложен литературным языком с использованием современной терминологической терминологии. Могут быть допущены 2-3 неточности или незначительные опечатки, и исправленные студентом с помощью преподавателя.

Отметка «УДОВЛЕТВОРИТЕЛЬНО» - даны недостаточно полный и недостаточно развернутый ответы. Логика и последовательность изложения имеют нарушения. Допущены ошибки в раскрытии понятий, употреблении терминов. В ответе отсутствуют выводы. Умение раскрыть значение обобщенных знаний не показано. Речевое оформление требует поправок, коррекции.

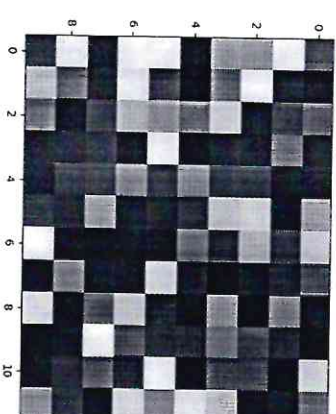
Отметка «НЕУДОВЛЕТВОРИТЕЛЬНО» - ответ представляет собой разрозненные знания с существенными ошибками по вопросам. Присутствуют фрагментарность, неполнота изложения. Отсутствуют выводы, конкретизация и доказательность изложения. Речь неграмотная, тисготит чуждая терминология не используется. Дополнительные и уточняющие вопросы преподавателя не приводят к коррекции ответа студента. Или ответ на вопрос полностью отсутствует, или отказ от ответа.

ЗАДАНИЯ

ДЛЯ ОЦЕНКИ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ

Задания для оценки сформированности компетенции ОПК-3:

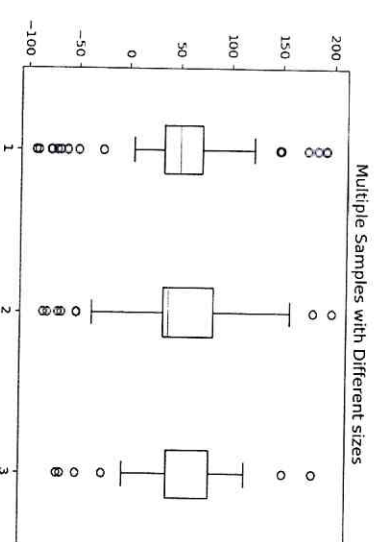
1. Какого типа следующий график?



1. bar
2. hist
3. box
4. scatter
5. heatmap

Ответ: 5

2. Какого типа следующий график?



1. bar
2. hist
3. box
4. scatter
5. heatmap

Ответ: 3

3. Что такое NumPy?

1. Python библиотека для работы с большими многомерными массивами и матрицами, имеющая большой набор математических функций для операций с этими массивами.
2. Библиотека Python для работы с числами, строками и другими типами данных.
3. Функция Python для анализа больших данных.

Ответ: 1

4. У нас есть следующий объект DataFrame:

```
df = pd.DataFrame({
    'country': ['Kazakhstan', 'Russia', 'Belarus', 'Ukraine'],
    'population': [17.04, 143.5, 9.5, 45.5],
    'square': [2724902, 17125191, 207600, 603628]
})
```

При этом мы заменим индексы следующим способом:

```
df.index = ['KZ', 'RU', 'BU', 'UA']
```

Что нам надо сделать, чтобы добавить наименование индексов?

1. df.index.name = 'Country Code'
2. df.add(name='Country Code')
3. df.append.name('Country Code')

Ответ: 1

5. У нас есть файл с данными в формате "csv". Какой тип данных лучше всего подойдет, если мы хотим работать только с единственной колонкой с данными?

1. Series
2. DataFrame
3. Словарь
4. Кортеж

Ответ: 1

6. Что будет содержать объект Series?

```
my_series = pd.Series([5, 6, 7, 8, 9, 10])
```

- 1) 0 5
1 6
2 7
3 8
4 9
5 10
- 2) 1 5
2 6
3 7
4 8
5 9
6 10
- 3) 5
6
7
8
9
10

Ответ: 1

7. У нас имеется файл с данными в формате "csv". Какой тип данных лучше всего подойдет, если мы хотим работать со всей информацией, хранящейся в данном файле?

1. Series
2. DataFrame
3. Словарь
4. Список

Ответ: 2

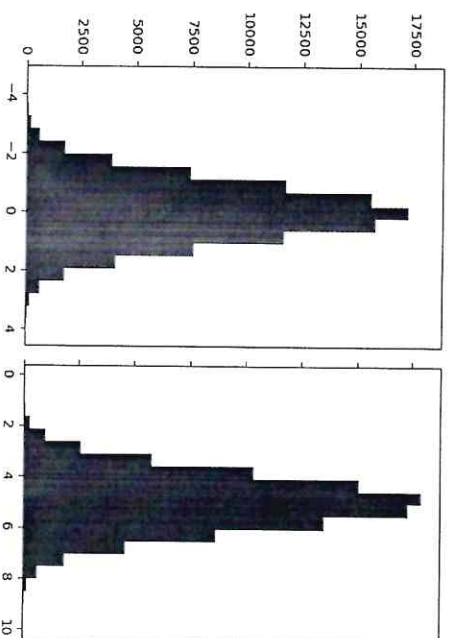
8. Что из себя представляет объект DataFrame?

1. Объект библиотеки Pandas, представляющий из себя табличную структуру данных.
2. Объект библиотеки Pandas, являющуюся словарем.

3. Python объект, описывающий формат данных типа "Дата"

Ответ: 1

9. Какого типа следующий график?



1. bar
2. hist
3. box
4. scatter
5. heatmap

Ответ: 2

10. С помощью какого метода можно построить сводную таблицу в Pandas?

1. .pivot_table
2. .groupby
3. read_csv

Ответ: 1

11. Можно ли представить объект Series следующим образом?

`my_series = pd.Series({'a': 5, 'b': 6, 'c': 7, 'd': 8})`

1. Да, можно
2. Нет, нельзя

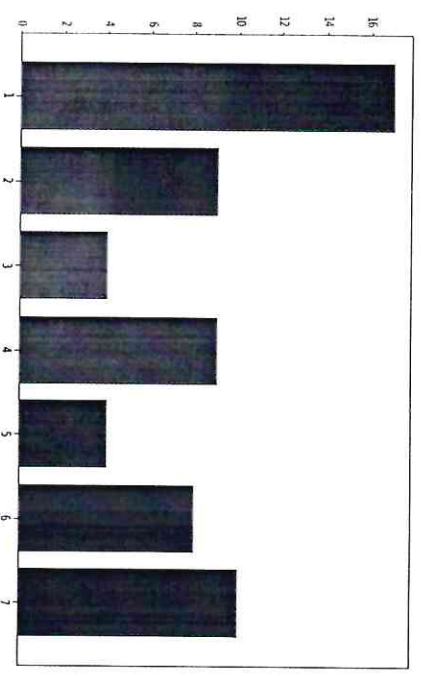
Ответ: 1

12. Что такое Pandas?

1. Библиотека Python для анализа данных.
2. Библиотека Python для работы с графикой.
3. Встроенная Python функция для анализа больших данных.

Ответ: 1

13. Какого типа следующий график?



1. bar
2. hist
3. box
4. scatter
5. heatmap

Ответ: 1

14. Что из себя представляет объект Series?

1. Структура из библиотек Pandas, похожая на одностолбчатый массив с использованием ассоциированных меток.
2. Структура из библиотек Pandas, похожая на табличную структуру данных.
3. Объект библиотеки NumPy, являющийся Python-овским списком

Ответ: 1

15. Агрегирование данных это...
16. Визуализация данных это...
17. Характеристики Nlp...
18. Характеристики Matplotlib...
19. Характеристики Pandas...
20. Основные объекты для анализа данных...
21. История и скусственн ого интеллекта...
22. Методы машинного обучения...
23. Задача классификации:
24. Регрессия это...
25. Задача кластеризации это...

Критерии оценки результатов:

- оценка «отлично» выставл яется студенту, если он отвечает верно на 80-100% вопросов.
- оценка «хорошо» выставл яется студенту, если он отвечает верно на 70-79% вопросов.
- оценка «удовлетворительно» выставл яется студенту, если он отвечает верно на 60-69% вопросов.
- оценка «неудовлетворительно» выставл яется студенту, если он не освоил материал темы, дает менее 60% правильных ответов.

МАТРИЦА СООТВЕТСТВИЯ КРИТЕРИЕВ ОЦЕНКИ УРОВНЮ С ФОРМИРОВАНОСТИ КОМПЕТЕНЦИЙ

Критерии оценки	Уровень сформированности компетенций
Оценка по пятибалльной системе	
«Отлично»	«Высокий уровень»
«Хорошо»	«Повышенный уровень»
«Удовлетворительно»	«Пороговый уровень»
«Неудовлетворительно»	«Не достаточный»
Оценка по системе «зачет – не зачет»	
«Зачтено»	«Достаточный»
«Не зачтено»	«Не достаточный»

Методические материалы, определяющие процедуру оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

1. Положение «О балльно-рейтинговой системе аттестации студентов»: СМК ПНД 08-01-2022, введено приказом от 28.09.2011 №371-О (<http://nsu.edu.ru/file/403>: режим доступа свободный);
2. Положение «О проведении текущего контроля и промежуточной аттестации обучающихся в ФГБОУ ВО Новосибирский ГАУ»: СМК ПНД 77-01-2022, введено в действие приказом от 03.08.2015 №268а-О (<http://nsu.edu.ru/file/104821>: режим доступа свободный).