

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Кислова Наталья Николаевна

Должность: Проректор по УЧ и качеству образования

Дата подписания: 03.06.2024

Уникальный программный ключ:

52802513f5b14a975b3e9b13008093d5726b159bf6064f865ae65b96a966c035

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

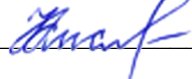
Федеральное государственное бюджетное образовательное учреждение высшего образования

«Самарский государственный социально-педагогический университет»

Кафедра информатики, прикладной математики и методики их преподавания

Утверждаю

Проректор по учебно-методической
работе и качеству образования

 Н.Н. Кислова

Макарова Елена Леонидовна

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

для проведения промежуточной аттестации по дисциплине
«Введение в анализ данных»

Направление подготовки:

09.03.03 Прикладная информатика

Направленность (профиль):

«Корпоративные информационные системы»

Квалификация выпускника

Бакалавр

Рассмотрено

Протокол №11 от 25.06.2024 г.

Заседания кафедры информатики, прикладной
математики и методики их преподавания

Одобрено

Начальник Управления образовательных
программ



Н.А. Доманина

Пояснительная записка

Фонд оценочных средств (далее – ФОС) для промежуточной аттестации по дисциплине «Введение в анализ данных государственных органов» разработан в соответствии с ФГОС ВО по направлению подготовки 09.03.03 Прикладная информатика, утвержденного приказом Министерства образования и науки Российской Федерации от 19 сентября 2017г. №922, основной профессиональной образовательной программой «Корпоративные информационные системы» с учетом требований профессионального стандарта 06.015 «Специалист по информационным системам», утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 18 ноября 2014 г. №896н (зарегистрирован Министерством юстиции Российской Федерации 24 декабря 2014 г., регистрационный №35361), с изменением, внесенным приказом Министерством труда и социальной защиты Российской Федерации от 12 декабря 2016 г. №727н (зарегистрирован Министерством юстиции Российской Федерации 13 января 2017 г., регистрационный №45230).

Цель ФОС для промежуточной аттестации – установление уровня сформированности части компетенций УК-2.

Задачи ФОС для промежуточной аттестации – контроль качества и уровня достижения образовательных результатов по формируемым в соответствии с учебным планом компетенциям:

УК-2. Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений

УК-2.1: Формулирует совокупность взаимосвязанных задач в рамках поставленной цели работы, обеспечивающих ее достижение. Определяет ожидаемые результаты решения поставленных задач

Знает: Формулирует совокупность взаимосвязанных задач в рамках поставленной проблемы анализа больших данных и подбирает необходимый инструментарий (библиотеки Python)

УК-2.2. Проектирует решение конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений

Умеет: проектировать решение конкретной задачи анализа данных, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений.

УК-2.3 Качественно решает конкретные задачи (исследования, проекта, деятельности) за установленное время

Владеет: навыками качественно решать конкретные задачи анализа данных, в том числе их сбора за установленное время;

УК-2.4. Публично представляет результаты решения задач исследования, проекта, деятельности

Владеет: навыками публичного представления результатов решения задач исследования, проекта, деятельности в области анализа данных.

Требование к процедуре оценки:

Помещение: компьютерный класс.

Оборудование: персональные компьютеры с выходом в интернет.

Инструменты: Visual Studio Code/Idle, предустановленные библиотеки NumPy, Pandas, Matplotlib.

Расходные материалы: бумага, ручка.

Доступ к дополнительным справочным материалам: не предусмотрен.

Нормы времени: 90 мин.

Комплект оценочных средств для проведения промежуточной аттестации

Проверяемая компетенция:

Универсальная компетенция УК-2.

Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений

Проверяемые индикаторы:

УК-2.1: Формулирует совокупность взаимосвязанных задач в рамках поставленной цели работы, обеспечивающих ее достижение. Определяет ожидаемые результаты решения поставленных задач

УК-2.2. Проектирует решение конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений

УК-2.3 Качественно решает конкретные задачи (исследования, проекта, деятельности) за установленное время

УК-2.4. Публично представляет результаты решения задач исследования, проекта, деятельности

Проверяемые образовательные результаты:

Знает: Формулирует совокупность взаимосвязанных задач в рамках поставленной проблемы анализа больших данных и подбирает необходимый инструментарий (библиотеки Python)

Умеет: проектировать решение конкретной задачи анализа данных, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений.

Владеет: навыками качественно решать конкретные задачи анализа данных, в том числе их сбора за установленное время; навыками публичного представления результатов решения задач исследования, проекта, деятельности в области анализа данных.

Задание 1.

Тип (форма) задания: тест.

Содержание задания:

1. Отметьте неверное утверждение

- a) Python поддерживает только объектно-ориентированную парадигму программирования;
- b) Python – проект с открытым исходным кодом;
- c) Python – популярный язык веб-разработки;
- d) Python – популярный язык для задач искусственного интеллекта;
- e) Python поддерживает функциональную парадигму программирования.

2. В data science используются следующие библиотеки Python

- a) NumPy;
- b) SciPy;
- c) StatModels;
- d) math;
- e) random.

3. В интерактивном режиме диалог с пользователем имеет вид:

```
In [1]: 45 + 72
```

```
Out[1]: 117
```

Какое инструментальное средство используется

- a) Интерпретатор IPython;
- b) IDLE Python;
- c) Jupiter Notebook;
- d) Wing;
- e) Eclipse?

4. Очистка данных — ...

- a) комплекс методов и процедур, направленных на устранение причин, мешающих корректной обработке: аномалий, пропусков, дубликатов, противоречий, шумов и т.д.
- b) процесс дополнения данных некоторой информацией, позволяющей повысить эффективность развязки аналитических задач
- c) объект, содержащий структурированные данные, которые могут оказаться полезными для развязки аналитического задачи
- d) комплекс методов и процедур, направленных на извлечение данных из различных источников, обеспечение необходимого уровня их информативности и качества, преобразования в единый формат, в котором они могут быть загружены в хранилище данных или аналитическую систему

5. Какие способы позволяют отобразить 5 первых строк датафрейма dataf:

- a) dataf.head(5); dataf.iloc[0:5]; df.tail(7);
- b) dataf.head(5); dataf.iloc[0:5]; dataf.head(5);
- c) dataf.head(5); dataf.head(5); df.tail(7);
- d) dataf.iloc[0:5]; dataf.head(5); df.tail(7).

6. К библиотекам анализа данных в Python НЕ относятся ...

- a) tkinter
- b) Pandas;
- c) NumPy;
- d) Matplotlib.

7. Общие затраты на производство данных в настоящее время достигают

- a) 3% -10% от мирового производства энергии;
- b) до 1% мирового производства энергии;
- c) свыше 30% мирового производства энергии;

- d) незначительны в мировом потреблении электроэнергии;
- e) нет верного ответа.

8. К четырем основным характеристикам больших данных, обычно называемым «четырьмя V», не относятся:

- a) актуальность;
- b) объем;
- c) скорость;
- d) разнообразие;
- e) достоверность.

9. GPS-приложение Google Waze анализирует пользовательские данные для определения оптимального маршрута к месту назначения за минимальное время (аналог в России – Яндекс.Пробки). При этом пользователи также являются источником данных о различных препятствиях на дорогах. Такое приложение использует следующие сквозные технологии:

- a) синтез речи;
- b) NLP (обработка естественного языка);
- c) большие данные (динамическую визуализацию);
- d) IoT;
- e) машинное и глубокое обучение.

10. Команда `print(56//5+12%5)` в Python в качестве результата выведет

- a) 13;
- b) сообщение об ошибке;
- c) 13.2;
- d) `56//5+12%5`;
- e) 8.

11. Каким образом можно извлечь корень квадратный в Python из заданного числа 7

- a) `7**(1/2)`;
- b) `7**0.5`;
- c) `sqrt(7)`;
- d) `import math;`
`print (math.sqrt(7))`;
- e) `sqrt(7)`?

12. Строка в Python – последовательность символов, заключенная в

- a) апострофы;
- b) двойные кавычки;
- c) круглые скобки;
- d) квадратные скобки;
- e) фигурные скобки.

13. Функция `input` запрашивает данные у пользователя и получает их (в версиях 3.* Python) как

- a) строку;
- b) число;
- c) зависит от конкретного значения;
- d) список;
- e) указатель на объект.

14. Условие «значение x принадлежит отрезку [a;b]» на языке Python записывается как

- a) `a<=x<=b`;
- b) `a<x<b`;
- c) `(a<=x) and (x<=b)`;
- d) `(a<=x) or (x<=b)`;
- e) `(a<=x) and (x<=b)`.

15. Для нахождения минимума среди значений 4 переменных a,b,c,d в Python можно использовать

- a) `min(a,b,c,d)`;

- b) `min(min(a,b),min(c,d))`;
- c) `min(a,min(b,c,d))`;
- d) написать программу с применением условных операторов;
- e) `a+b+c+d-max(a,b,c,d)-max(a,b,c)`.

16. Если вместо условия (логического выражения) после ключевого слова `if` в Python указать число 7, то

- a) число будет интерпретировано как `True`;
- b) число будет интерпретировано как `False`;
- c) будет выведено сообщение об ошибке;
- d) результат зависит от используемой среды разработки IDE;
- e) результат зависит от настроек редактора.

17. Если значение переменной `grade = 56`, то программный код выведет

```
if grade >= 90:
```

```
...: print('A')
...: elif grade >= 80:
...: print('B')
...: elif grade >= 70:
...: print('C')
...: elif grade >= 60:
...: print('D')
...: else:
...: print('F')
```

- a) F;
- b) D;
- c) C;
- d) B;
- e) A.

18. Дан программный код вычисления среднего балла в заданном наборе оценок. Какая строка пропущена (на ее месте знаки `???`)

```
1 # class_average.py
2 """Вычисление средней оценки с повторением, управляемым последовательностью."""
3
4 # Фаза инициализации
5 total = 0 # Сумма оценок
6 grade_counter = 0
7 grades = [98, 76, 71, 87, 83, 90, 57, 79, 82, 94] # Список из 10 оценок
8
9 # Фаза обработки
10 for grade in grades:
11     ????? # Прибавить текущую оценку к накапливаемой сумме
12     grade_counter += 1 # Еще одна оценка была обработана
13
14 # Завершающая фаза
15 average = total / grade_counter
16 print(f'Class average is {average}')
```

- a) `total += grade`;
- b) `total =total+grade`;
- c) `total *= grade`;
- d) `total =sum(grade)`;
- e) `total =avg(grade)?`

19. Что будет возвращено при выполнении следующего кода в IPython

```
In [1]: import numpy as np
In [2]: numbers = np.array([2, 3, 5, 7, 11])
In [3]: type(numbers)
```

- a) Out[3]: numpy.ndarray;
- b) Out[3]: np.ndarray;
- c) Out[3]: ndarray;
- d) Out[3]: array;
- e) Out[3]: list?

20. Для получения количества измерений массива numpy используется атрибут

- a) .ndim;
- b) .shape;
- c) .size;
- d) .itemsize;
- e) не атрибут, а функция len(array).

21. Массив `grades1=np.array([120,130,140,135])`, `grades2= np.array([125,126,127,126])`. Не выполняя вычислений, оцените что вернет команда `grades1.var()<grades2.var()`

- a) False;
- b) True;
- c) сообщение об ошибке;
- d) 135;
- e) 126.

22. К какой категории универсальных функций NumPy относится `trunc`

- a) Функции с плавающей точкой;
- b) Функции сравнения;
- c) Поразрядные функции;
- d) Тригонометрические;
- e) Математические?

23. Какое из утверждений относительно представлений и глубокого копирования массивов numpy является неверным

- a) Глубокое копирование можно применять только к небольшим по размеру массивам;
- b) Для создания представлений используется метод `view()`;
- c) Для глубокого копирования используется метод `copy()`;
- d) Изменения исходного массива отражаются на тех представлениях, которые содержат модифицированные элементы;
- e) Для глубокого копирования некоторых объектов Python используется функция `deepcopy()`?

24. В библиотеке Pandas для работы с наборами данных в основном используются

- a) Series;
- b) DataFrame;
- c) `np.array`;
- d) списки;
- e) кортежи.

25. Метод `describe()` для коллекции Series (pandas) возвращает

- a) основные описательные статистики для набора данных;
- b) только среднее значение;
- c) только минимум и максимум;
- d) зависит от установленных параметров;
- e) количество элементов в коллекции.

26. Для того, чтобы очистить строку от пробелов слева и справа используется

- a) метод `strip()`;
- b) метод `lstrip()`;
- c) метод `rstrip()`;
- d) метод `trunc()`;
- e) срезы.

27. Для разбиения строк на части по разделителю можно использовать методы

- a) split();
- b) rsplit();
- c) partition();
- d) splitlines();
- e) rpartition().

28. Что из перечисленного является квантификатором

- a) {7};
- b) \d;
- c) \s;
- d) \w;
- e) [aeiou]?

30. Первичная обработка (подготовка данных) включает

- a) очистку данных (исключение выбросов, устранение дубликатов,...);
- b) преобразование данных в форматы, оптимальные для хранения и обработки;
- c) вычисление показателей описательной статистики;
- d) подготовку отчета;
- e) организацию резервного хранения данных.

30. Почему часть специалистов data science считает некорректным подход со вставкой «разумных значений (средних, часто встречающихся) и т.п.»

- a) подстановка часто превращается в подтасовку;
- b) не всегда можно выбрать разумное значение;
- c) нельзя исключать данные, противоречащие гипотезе;
- d) это не предусмотрено некоторыми пакетами анализа данных;
- e) это приводит к слишком большой погрешности результатов?

Правильные ответы к заданию 1

1	a	11	a,b,d	21	a
2	a,b,c	12	a,b	22	a
3	a	13	a	23	a
4	a	14	a,c	24	a,b
5	b	15	a,b,c,d	25	a
6	a	16	a	26	a
7	a	17	a	27	a,b,c,d,e
8	a	18	a,b	28	a
9	a,b,c,d,e	19	a	29	a,b
10	a	20	a	30	a,b,c

Оценочный лист к заданию 1.

Критерий	Максимальное количество баллов
Вопрос 1-30	1

Проверяемые индикаторы:

УК-2.1: Формулирует совокупность взаимосвязанных задач в рамках поставленной цели работы, обеспечивающих ее достижение. Определяет ожидаемые результаты решения поставленных задач

УК-2.2. Проектирует решение конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений

УК-2.3 Качественно решает конкретные задачи (исследования, проекта, деятельности) за установленное время

УК-2.4. Публично представляет результаты решения задач исследования, проекта, деятельности

Проверяемые образовательные результаты:

Знает: Формулирует совокупность взаимосвязанных задач в рамках поставленной проблемы анализа больших данных и подбирает необходимый инструментарий (библиотеки Python)

Умеет: проектировать решение конкретной задачи анализа данных, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений.

Владеет: навыками качественно решать конкретные задачи анализа данных, в том числе их сбора за установленное время; навыками публичного представления результатов решения задач исследования, проекта, деятельности в области анализа данных.

Задание 2.

Тип (форма) задания: практическая задача.

Содержание задания: скачайте с портала «Открытые данные России» Данные Информация об обращениях граждан (сведения о тематике и результатах рассмотрения) <https://data.gov.ru/opendata/7710349494-og>. Подготовьте данные к использованию Средствами языка программирования Python определите наименование вопроса, по которому поступило максимальное количество обращений в Министерство в 4 квартале, и вопросы, по которым поступило минимальное количество обращений. Проведите визуализацию полученных результатов.

Оценочный лист к заданию 2.

Указания по оцениванию	Индикатор УК-2	Баллы
Данные скачены; сформирован датасет;	УК2.1	5
Данные очищены, из выборки исключена последняя строка, содержащая суммарное количество обращений;	УК2.2	5
Выведены максимальное количество обращений в Министерство в 4 квартале, и вопросы, по которым поступило минимальное количество обращений.	УК2.3	5
Представлен дашборд по данным	УК2.4	15
Максимальное число баллов за задание		30

Методические материалы, определяющие процедуру и критерии оценивания сформированности компетенций при проведении промежуточной аттестации

Код контролируемой компетенции (индикаторы)	Наименование оценочного средства	Максимальное количество баллов	Всего баллов	Уровень освоения компетенции (в баллах)		
				Пороговый (56-70%)	Продвинутый (71-85%)	Высокий (86-100%)
УК-2.1	Задание 1	30	60	36-41	42-51	52-60
УК-2.2						
УК-2.3						
УК-2.4						
УК-2.1	Задание 2	5				
УК-2.2	Задание 2	5				
УК-2.3	Задание 2	5				
УК-2.4	Задание 2	15				

Полученное число баллов выставляется в графу «Промежуточная аттестация» балльно-рейтинговой карты дисциплины.