

Демонстрационные задания аттестационных испытаний перевода на ОП «Инженерия данных»

Время выполнения заданий – 120 минут.

Разрешено пользоваться:

- калькулятором
- черновиком
- средой программирования из белого списка: Jupyter Notebook, Jupyter Lab, PyCharm, Thonny, Sublime Text, Python Shell, Wing Personal 7.2, Atom, Visual Studio Code, IDLE Shell, DataSpell, JupyterLite
- документациями библиотек pandas (<https://pandas.pydata.org/>), numpy (<https://numpy.org/doc/>), matplotlib (<https://matplotlib.org/stable/index.html>)

1. Пусть $f(x) = x^3 + x^2 + 3x + 7$. При каком наименьшем x производная этой функции равна 4?

Ответ: -1

2. В книжном шкафу лежат 12 томов Толстого, 10 томов Пушкина, 7 томов Чехова и 5 томов Достоевского. Сколькими способами можно выбрать 2 книги разных авторов?

Ответ: 419

3. В классе 33 ученика. На уроке им задали три задачи. Каждую задачу решили по 14 учеников. Каждую пару задач решили по 4 ученика. Кроме того, один ученик решил все три задачи. Сколько учеников не решило ни одной задачи?

Ответ: 2

4. При каком целом a имеет место равенство

$$\int_0^a (x^2 - x + 1) dx = \frac{15}{2} ?$$

Ответ: 3

5. Сколько существует пятизначных чисел, в которых есть хотя бы одна чётная цифра?

Ответ: 86875

6. Известно, что определитель матрицы

$$\begin{pmatrix} 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 3 \\ x & 10 & 2 & 1 \\ 5 & 17 & 6 & y \end{pmatrix}$$

равен 20. Найдите значение выражения $y + 9x$

Ответ: 13

7. В баскетбольной команде 9 человек. Для игры тренеру нужно расставить 5 игроков из команды на 5 разных позиций (остальные игроки будут сидеть на скамейке запасных). Сколькими способами он может это сделать?

Ответ: 15120

8. Решите систему линейных уравнений. В качестве ответа укажите значение выражения $\frac{p_1+p_2}{p_3+p_4}$, где (p_1, p_2, p_3, p_4) — произвольное ненулевое решение системы.

$$\begin{cases} 2x_1 + x_2 - x_3 + x_4 = 0; \\ x_1 + x_2 + 2x_3 + 3x_4 = 0; \\ x_1 - 4x_2 + x_3 - 4x_4 = 0. \end{cases}$$

Ответ: -6

9. В классе 6 девочек и 9 мальчиков. Сколькими способами можно выбрать команду из 3 девочек и 3 мальчиков?

Ответ: 1680

10. В арифметическом линейном пространстве \mathbb{R}^4 даны 4 вектора $(8,0,4,\lambda)$, $(1,0,3,0)$, $(6,2,-8,7)$, $(-3,0,\lambda-1,-1)$. Укажите наименьшее значение параметра λ , при котором данные векторы НЕ образуют базис этого линейного пространства.

Ответ: -10

11. Какая из следующих функций является более «быстрой» (имеет меньшую асимптотическую сложность)?

1. $O(n^2)$
2. $O(\log \log n)$
3. $O(\log n)$
4. $O(n)$
5. $O(1)$

Ответ: 5

12. Квадратичный алгоритм с временем обработки $T(n) = cn^2$ тратит $T(N)$ секунд на обработку N элементов данных. Сколько времени будет потрачено на обработку $n = 5000$ элементов данных, при условии, что $N = 100$ и $T(N) = 1ms$?

Ответ: 2500

13. Пусть есть список из элементов $[25, 45, 87, 21, 18, 49, 13, 115, 83, 65]$ и искомое значение $x = 83$. Какое количество операций сравнения нужно совершить для линейного поиска элемента x в исходном списке?

1. 1
2. 7

3. 8
4. 9
5. 10

Ответ: 4

14. Какова временная сложность сортировки вставками в лучшем случае?

1. $O(n^2)$
2. $O(\log \log n)$
3. $O(\log n)$
4. $O(n)$
5. $O(1)$

Ответ: 4

15. Выберите структуру данных, работа с элементами которой организована по принципу FIFO (первый пришел - первый ушел)

1. Стек
2. Словарь
3. Список
4. Очередь

Ответ: 4

16. Что такое область видимости переменной?

1. Часть кода, где переменная была определена
2. Часть кода, где переменная была использована
3. Часть кода, где переменная была передана в функцию
4. Часть кода, где переменная была удалена

Ответ: 1

17. Что не относится к стандарту кодирования?

1. CP866
2. KOI8-R
3. UTF-8
4. TCP/IP

Ответ: 4

18. Как называется процесс поиска ошибок и их исправления в программе?

1. Отладка
2. Компиляция
3. Исполнение
4. Тестирование

Ответ: 1

19. Вам дан датасет titanic.csv со следующими параметрами:

- PassengerId: Уникальный индекс/номер строки.
- Survived: Признак, показывающий был ли спасен данный пассажир или нет.
- Pclass: Класс билета.
- Name: Имя пассажира с указанием титула.
- Sex: Пол пассажира.
- Age: Возраст пассажира.
- SibSp: Количество братьев/сестер или супругов, путешествующих с каждым пассажиром.
- Parch: Количество родителей детей
- Ticket: Номер билета.
- Fare: Сумма, которую заплатил пассажир за путешествие.
- Cabin: Номер каюты пассажира.
- Embarked: Порт отправления данного пассажира.

Сколько уникальных титулов пассажиров?

Ответ: 17

20. Продолжаем работать с датасетом titanic.csv. Какая доля выживших в зависимости от титула и наличия родственников на борту?

Объедините титулы, которые встречаются менее 10 раз, в категорию "Other".

Ответы расположите в порядке возрастания в формате:

Титул, наличие родственников (0, 1): доля выживших, округленная до 2 знаков после точки

Ответ:

```
Mr, 1: 0.15
Mr, 0: 0.17
Other, 0: 0.43
Other, 1: 0.45
Master, 0: 0.57
Miss, 0: 0.63
Miss, 1: 0.75
Mrs, 0: 0.77
Mrs, 1: 0.90
```