



ИНТЕЛЛЕКТУАЛЬНЫЙ
МЕГАПОЛИС

ЗАДАЧНИК



ИТ-класс

В МОСКОВСКОЙ ШКОЛЕ

**НАПРАВЛЕНИЕ
БОЛЬШИЕ ДАННЫЕ И
ТЕХНОЛОГИИ ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА**

ПРАКТИЧЕСКИЙ ЭТАП

**МОСКВА
2024**

Оглавление

Вариант 1	3
Вариант 2	10
Вариант 3	16
Вариант 4	22
Вариант 5	27
Вариант 6	32
Вариант 7	37
Вариант 8	42
Вариант 9	47
Вариант 10	54
Вариант 11	60
Вариант 12	66
Вариант 13	72
Вариант 14	77
Вариант 15	82
Вариант 16	87

Вариант 1

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_1

Задача 1.

В группе из 10 человек нужно выбрать 4 для участия в соревнованиях. Сколько различных способов существует для выбора команды?

В ответ запишите развернутое решение. Для обоснования воспользуйтесь символами: $C(n, k)$ - для обозначения сочетаний.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

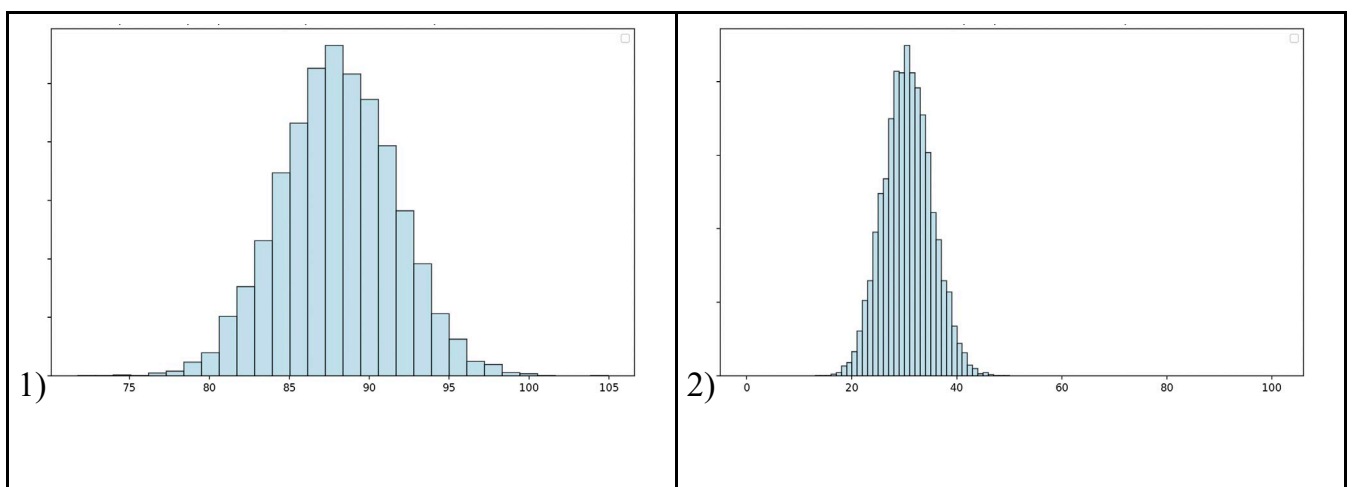
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

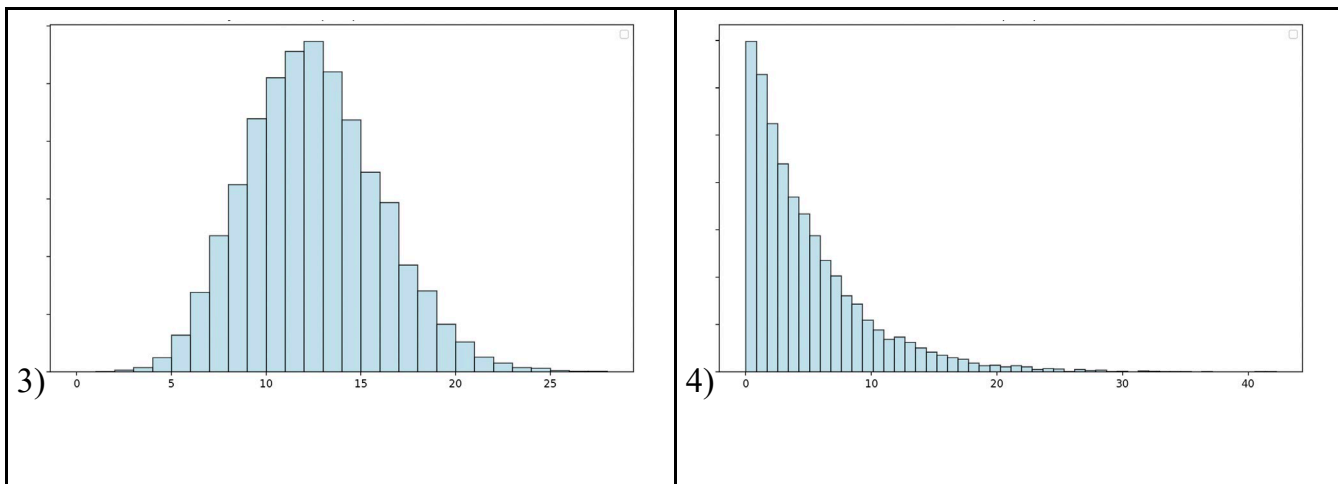
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Число заказов в интернет-магазине за час: Среднее количество заказов в интернет-магазине за час составляет 12 заказов.

Б) Распределение оценок производительности сотрудников: В крупной компании оценки производительности сотрудников могут варьироваться, с большинством оценок в среднем диапазоне и меньшим количеством высоких или низких оценок.

В) Распределение времени до следующего отказа в системе: Среднее время, которое проходит между двумя отказами в компьютерной системе или сети, составляет 5 минут.

Г) Распределение числа положительных отзывов на продукт из 100 опрошенных пользователей, при вероятности положительного отзыва 0,3.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая вычисляет сумму элементов списка целых чисел.

Формат ввода:

Целые числа, разделенные пробелом.

Формат вывода: Сумма элементов списка: <сумма>

Пример:

Ввод:

1 2 3 4 5

Вывод:

Сумма элементов списка: 15

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма (ГГГГ-ММ-ДД), position - должность, age - возраст сотрудника.

Задача: найти сотрудников, которые были наняты после 1 января 2020 года.

Формат вывода: Имена сотрудников, дата найма

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (sports_competition.csv), содержащий информацию о спортивных соревнованиях. Этот набор данных включает следующие столбцы:

1. Название соревнования
2. Участник
3. Страна
4. Год проведения
5. Вид спорта
6. Количество набранных очков
7. Количество медалей

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите топ-3 спортсменов, набравших наибольшее количество очков за последние пять лет (с 2019 года до 2024 года).
2. Постройте круговую диаграмму, показывающую распределение медалей среди этих трех спортсменов.

Формат вывода:

Топ-3 спортсменов по количеству набранных очков:

<Имя спортсмена 1> (<Страна>) - количество очков: X

<Имя спортсмена 2> (<Страна>) - количество очков: Y

<Имя спортсмена 3> (<Страна>) - количество очков: Z

Решение задачи и полученные файлы сохраните в папке **задача_5**.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	Диаграмма, полученная в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Вам дан код для линейной регрессии. В данном коде есть ошибка - найдите ее и исправьте.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LinearRegression
import numpy as np

# Генерация неправильных данных
X = np.array([[1], [2], [3], [4], [5]])
y = np.array([1.5, 1.7, 1.9, 2.1])

# Обучение модели
model = LinearRegression()
model.fit(X, y)

# Предсказания
predictions = model.predict(X)
print(predictions)
```

Код (решение) запишете в файл `task_6.py` или `задача_6.txt`

Задача 7.

Вы работаете в компании, управляющей сетью кинотеатров. Необходимо спроектировать базу данных для хранения информации о фильмах, сеансах и билетах.

Условия задачи:

- * Фильмы: Каждый фильм имеет уникальный идентификатор, название, жанр и продолжительность.
- * Сеансы: Каждый сеанс привязан к одному фильму, имеет уникальный идентификатор, дату и время показа, зал.
- * Билеты: Каждый билет имеет уникальный идентификатор, привязан к сеансу и месту в зале, включает цену.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о фильмах, сеансах и билетах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения всех сеансов с указанием **названия фильма, жанра, времени показа и зала**. Для этого запрос должен объединять данные из таблиц **фильмов** и **сеансов**.

Примеры возможного заполнения таблиц:

Таблица Фильмы:

id	title	genre	duration
1	Inception	Sci-Fi	148
2	Titanic	Drama	195

3	The Matrix	Action	136
4	Avatar	Sci-Fi	162
5	The Godfather	Crime	175

Таблица Сеансы:

id	movie_id	screening_time	hall
1	1	2024-09-10 18:00:00	Hall 1
2	2	2024-09-10 20:00:00	Hall 2
3	3	2024-09-11 15:30:00	Hall 3
4	4	2024-09-11 19:00:00	Hall 1
5	5	2024-09-12 21:00:00	Hall 2

Таблица Билетов:

id	screening_id	seat	price
1	1	A1	10.50
2	1	A2	10.50
3	2	B3	12.00
4	2	B4	12.00
5	3	C5	11.00
6	4	D1	13.50
7	5	E2	14.00

Решение задачи и полученные файлы сохраните в папке задача_7
Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
--------------	-------------------------	-----------------	----------------

Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 2

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_2

Задача 1.

В компании 70% сотрудников - программисты, а 30% - дизайнеры. 80% программистов имеют высшее образование, а 50% дизайнеров - нет. Какова вероятность того, что случайно выбранный сотрудник имеет высшее образование?

В ответ запишите развернутое решение. Для обоснования воспользуйтесь символами: $P(A)$ - для обозначения вероятности.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

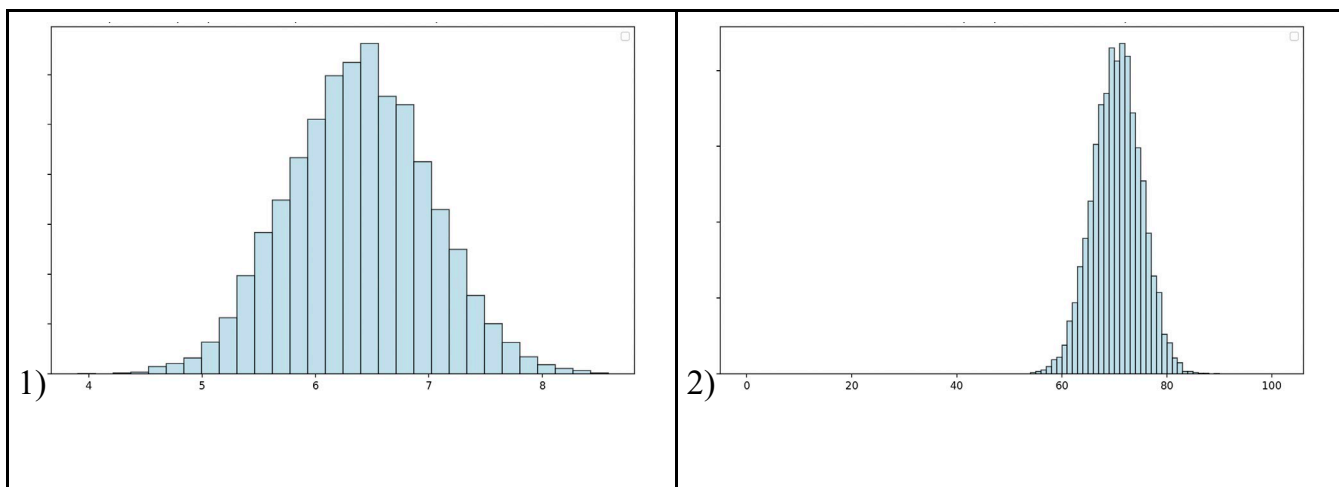
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

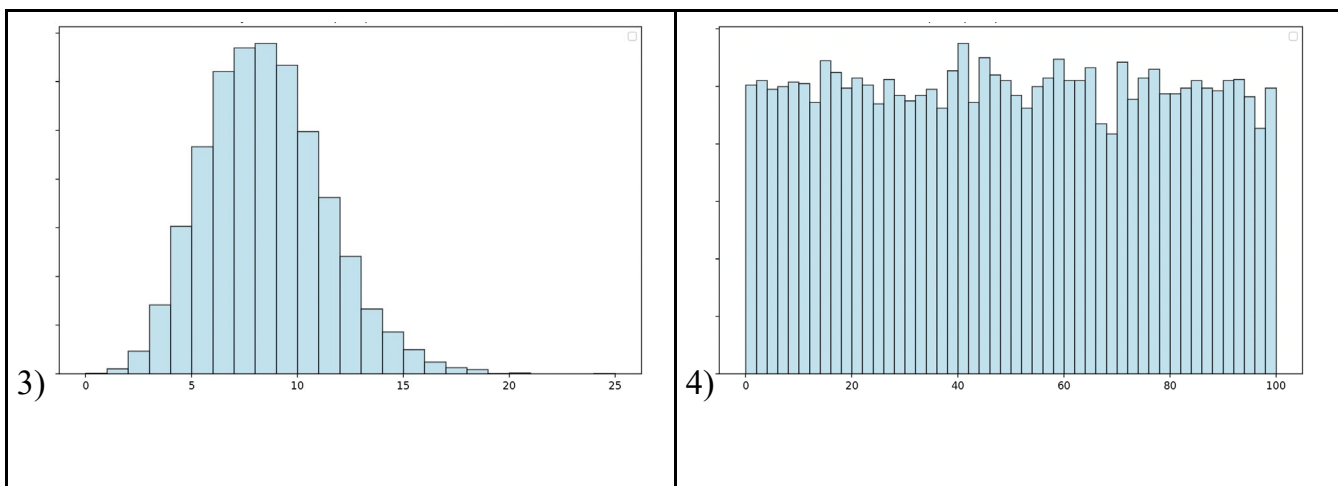
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Распределение скорости ветра в течение года на одной метеостанции: Ежегодное распределение скорости ветра на определенной метеостанции, будет варьироваться, с преобладанием средних значений и редкими крайними отклонениями.

Б) Число сбоев в работе серверов за месяц: Среднее количество сбоев в работе серверов за месяц составляет 8.

В) Распределение чисел на рулетке. Числа на колесе рулетки распределены равномерно, при этом каждый номер имеет равную вероятность выпадения.

Г) Распределение числа пациентов с положительной реакцией на лекарство из 100 испытуемых, при вероятности успеха 0,7.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая вычисляет сумму всех чисел от 1 до N (включительно), где N (целое число) вводится с клавиатуры.

Формат вывода: Сумма чисел от 1 до <N>: <сумма>.

Пример:

Ввод:

5

Вывод:

Сумма чисел от 1 до 5: 15.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма, position - должность, age - возраст сотрудника.

Задача: Отсортировать сотрудников по зарплате в порядке убывания.

Формат вывода: Имя сотрудника, зарплата.

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (sports_competition.csv), содержащий информацию о спортивных соревнованиях. Этот набор данных включает следующие столбцы:

1. Название соревнования
2. Участник
3. Страна
4. Год проведения
5. Вид спорта
6. Количество набранных очков
7. Количество медалей

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите вид спорта, в котором спортсмены набрали наибольшее количество очков за все время.
2. Постройте линейный график, показывающий изменение количества очков по годам проведения соревнований в этом виде спорта.

Формат вывода:

Вид спорта с наибольшим количеством набранных очков: <Название вида спорта>
- количество очков: <X>

Решение задачи и полученные файлы сохраните в папке задача_5.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код ниже можно использовать для реализации метода К-ближайших соседей.

Найдите ошибку в коде и исправьте её

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier

# Загружаем набор данных
iris = load_iris()
X, y = iris.data, iris.target

# Разделяем данные
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

# Создание модели KNN
knn = KNeighborsClassifier(n_neighbors=-3)
knn.fit(X_train, y_train)

# Оценка точности
accuracy = knn.score(X_test, y_test)
print("Точность:", accuracy)
```

Код (решение) запишете в файл task_6.py или задача_6.txt

Задача 7.

Вы разрабатываете базу данных для автосалона, который продает автомобили и ведет учет клиентов и продаж.

Условия задачи:

- * Машины: Автомобили имеют уникальный идентификатор, модель, производителя и цену.
- * Клиенты: Каждый клиент имеет уникальный идентификатор, имя и контактную информацию.
- * Продажа: Продажи автомобилей ведутся с указанием даты продажи и клиента, который купил автомобиль.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации автомобилях, клиентах и продажах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка всех продаж с указанием **модели автомобиля, производителя, имени клиента и даты продажи**. Для этого запрос должен объединять данные из таблиц **машин, клиентов и продаж**.

Решение задачи и полученные файлы сохраните в папке задача_7
Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 3

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_3

Задача 1.

Даны значения: 5, 8, 7, 10, 4. Какова медиана этих данных?

В ответ запишите развернутое решение.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

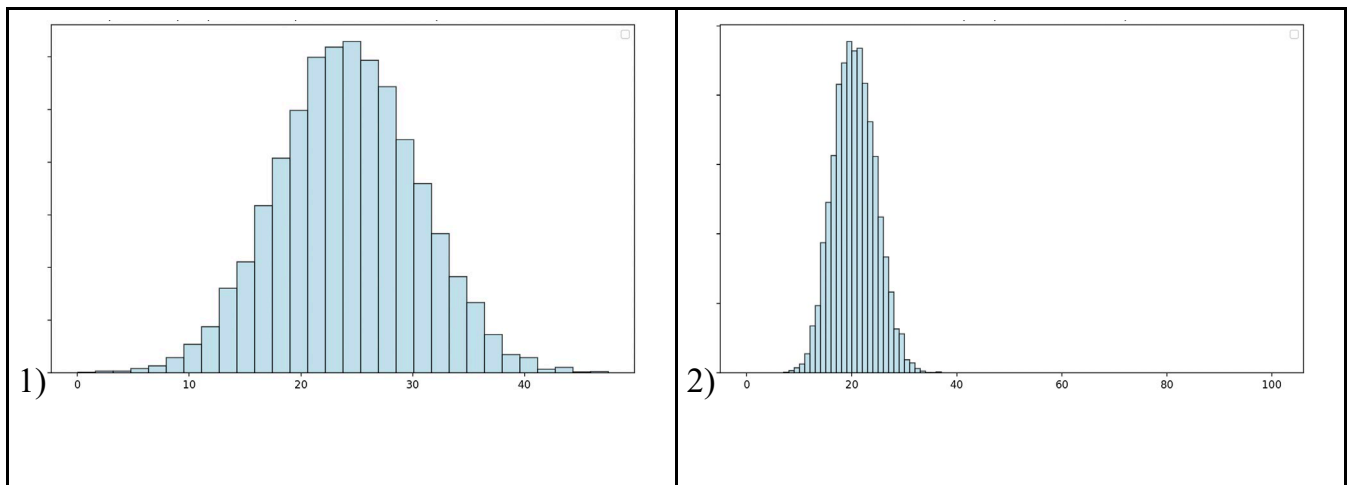
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

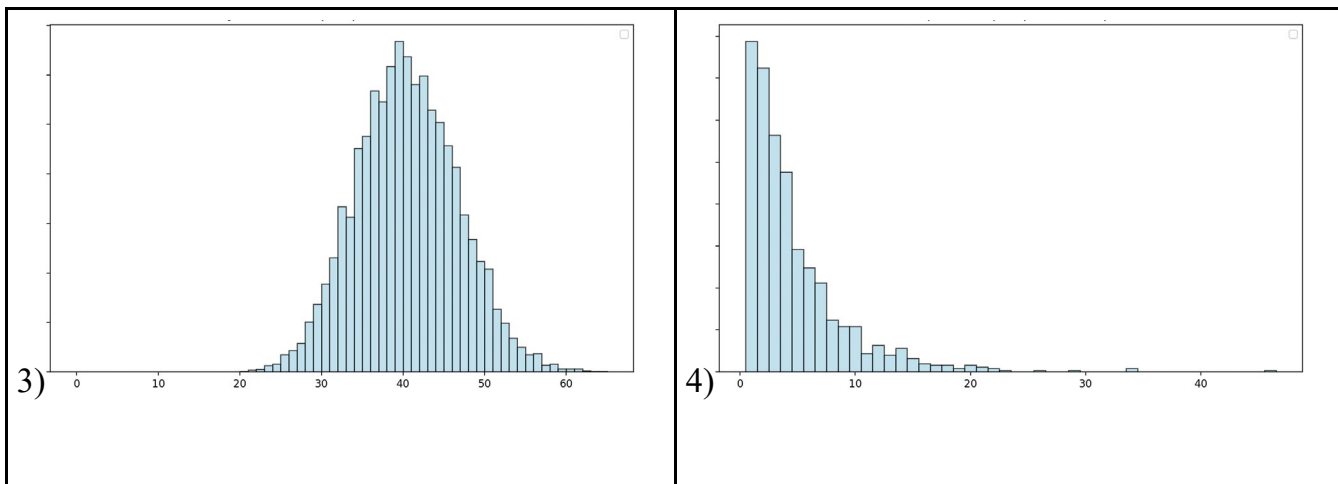
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Распределение числа успешных сборов пожертвований при обзвоне 100 человек, при вероятности согласия 0,2.

Б) Число писем, поступивших в почтовый ящик за день: Среднее количество писем, поступивших в почтовый ящик за день, равно 40.

В) Распределение описывает количество попыток на экзамене до первой успешной сдачи: Число попыток сдачи экзамена, необходимых для первой успешной сдачи.

Г) Распределение объемов продаж в розничной сети: В большой сети магазинов ежедневные объемы продаж на один магазин могут различаться, с преобладанием среднего уровня продаж.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая находит и выводит позицию первого вхождения подстроки `end` в строке. Если подстрока не найдена, выведите -1.

Формат вывода: Позиция первого вхождения подстроки `end` в строке <введенная строка>: <позиция>.

Пример:

Ввод:

abcdefend

Вывод:

Позиция первого вхождения подстроки end в строке abcdefend: 6.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма, position - должность, age - возраст сотрудника.

Задача: Найти среднюю зарплату всех сотрудников.

Вывод: Средняя зарплата

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (airport_flights.csv), содержащий информацию рейсах в аэропорту. Этот набор данных включает следующие столбцы:

1. Номер рейса
2. Авиакомпания
3. Пункт назначения
4. Время вылета
5. Время прилета
6. Количество пассажиров
7. Задержка (в минутах)

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Найдите рейсы, которые задерживались более чем на 1 час за последний месяц.
2. Постройте столбчатую диаграмму, отображающую количество таких рейсов по пунктам назначения.

Формат вывода:

Количество рейсов с задержкой более 1 часа: <X>

Решение задачи и полученные файлы сохраните в папке задача_5 .

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
--------------	----------------	-----------------	----------------

1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

В следующем коде производится оценка модели, но в нем есть несколько ошибок. Найдите и исправьте их.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import mean_squared_error

# Загрузка данных
iris = load_iris()
X, y = iris.data, iris.target

# Разделяем данные
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

# Обучаем модель
model = DecisionTreeClassifier()
model.fit(X_train, y_train)

# Предсказания
y_pred = model.predict(X_test)

# Оценка качества
mse = mean_squared_error(y_test, y_pred)
print("Средняя ошибка:", mse)
```

Код (решение) запишете в файл task_6.py или задача_6.txt

Задача 7.

Вы создаете базу данных для ресторана, чтобы хранить информацию о блюдах, клиентах и заказах.

Условия задачи:

1. Блюда имеют уникальный идентификатор, название и цену.
2. Клиенты ресторана имеют уникальный идентификатор, имя и контактную информацию.
3. Заказы включают несколько блюд, и каждый заказ имеет уникальный идентификатор, дату и общую сумму.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о блюдах, клиентах и заказах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка всех заказов с указанием **названия блюда, имени клиента и даты заказа**. Запрос должен объединять данные из таблиц **блюд, клиентов и заказов**, а также учитывать связь между заказом и заказанными блюдами.

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 4

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_4

Задача 1.

В коробке 5 красных, 3 зеленых и 2 синих шара. Если извлечь 2 шара, какова вероятность того, что оба шара будут красными?

В ответ запишите развернутое решение. Для обозначения числа сочетаний воспользуйтесь: $C(n, k)$, а для обозначения вероятности P .

Запишите свое решение в файл: задача_1.txt.

Задача 2.

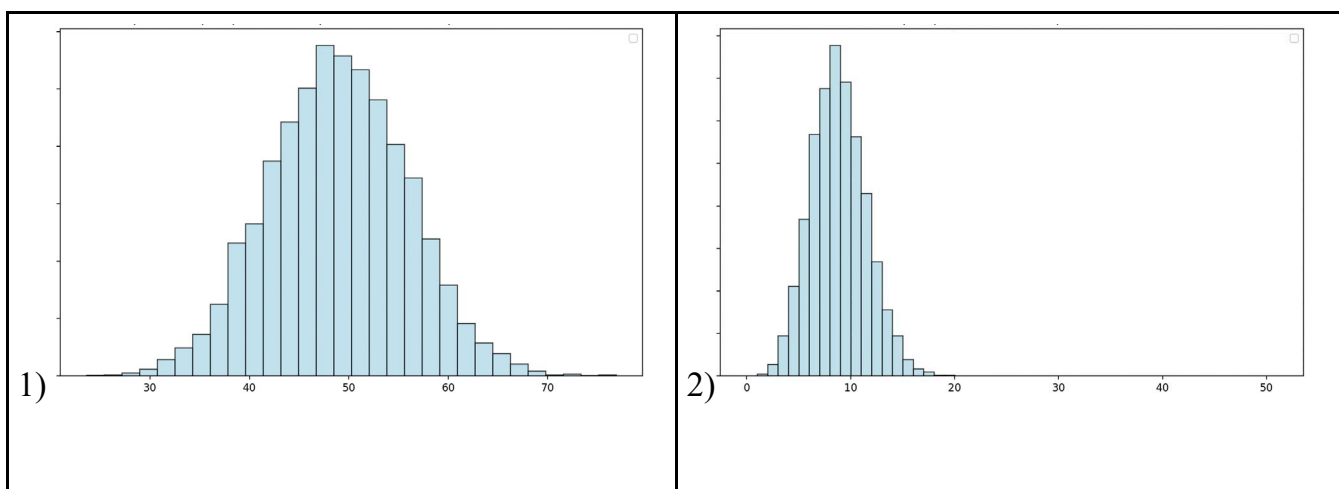
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

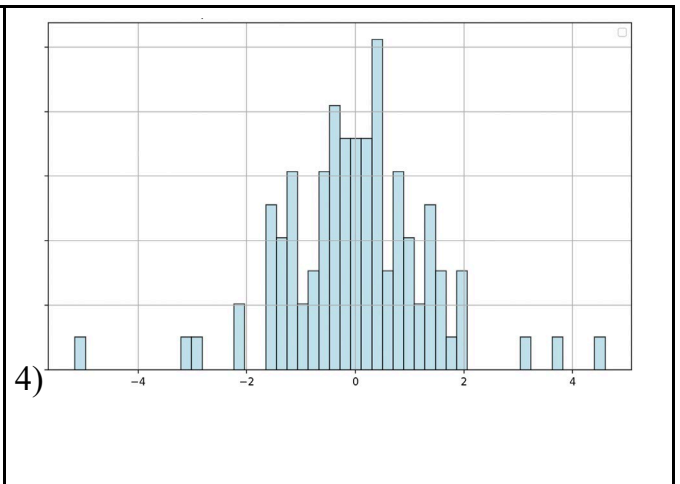
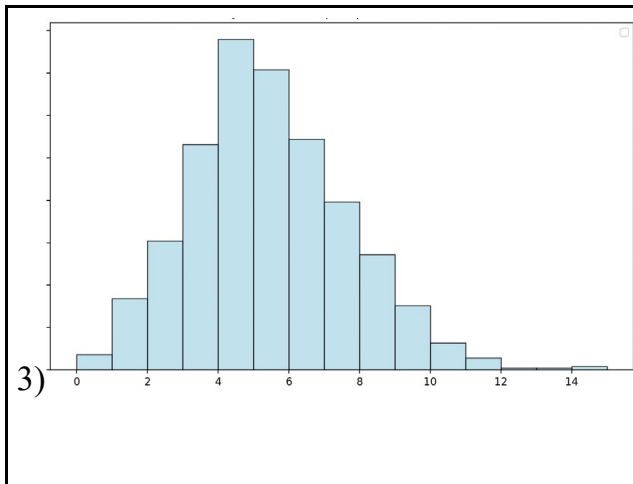
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Анализ результатов тестов на IQ в группе студентов: Применяется для оценки среднего IQ в группе студентов с небольшой выборкой

Б) Число дефектных изделий в партии из 1000 единиц: Среднее количество дефектных изделий в партии составляет 5.)

В) Распределение уровня физической активности среди населения: В большом городе уровень физической активности среди населения может быть разным, с большинством людей, ведущих умеренно активный образ жизни, и меньшинством, которое значительно менее или более активно.

Г) Распределение числа выпадений шестерки при 50 бросках кубика, где вероятность выпадения шестерки составляет $1/6$.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая заменяет все вхождения подстроки cat на dog в введенной строке.

Формат вывода: Строка после замены: <новая строка>.

Пример:

Ввод:

cat and catalog

Вывод:

Строка после замены: dog and dogalog.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма, position - должность, age - возраст сотрудника.

Задача: Вывести имена всех сотрудников, которые работают на позиции "Manager".

Вывод: имена сотрудников

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (airport_flights.csv), содержащий информацию рейсах в аэропорту. Этот набор данных включает следующие столбцы:

1. Номер рейса
2. Авиакомпания
3. Пункт назначения
4. Время вылета
5. Время прилета
6. Количество пассажиров
7. Задержка (в минутах)

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите топ-3 авиакомпаний по количеству пассажиров, перевезенных за последний месяц.
2. Постройте круговую диаграмму, показывающую распределение количества задержек среди этих авиакомпаний.

Формат вывода:

Топ-3 авиакомпаний по количеству пассажиров:

<Авиакомпания 1> - количество пассажиров: X

<Авиакомпания 2> - количество пассажиров: Y

<Авиакомпания 3> - количество пассажиров: Z

Решение задачи и полученные файлы сохраните в папке **задача_5**.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код создает и обучает модель линейной регрессии, но выводит ошибку при обучении. Найдите ошибку и исправьте ее.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LinearRegression
import numpy as np

X = np.array([[1, 2], [3, 4], [5, 6]])
y = np.array([3, 6, 9])

model = LinearRegression()
model.fit(X, y.reshape(-1, 1))
```

Код (решение) запишете в файл `task_6.py` или `задача_6.txt`

Задача 7.

Вы работаете в компании, управляющей складом. Требуется база данных для учёта товаров, поставщиков и заказов на склад.

Условия задачи:

- * Товары: Уникальный идентификатор, название, цена и количество.
- * Поставщики: Уникальный идентификатор, название компании и контактные данные.
- * Заказы: Уникальный идентификатор, номер транзакции, поставщика и товар.

Транзакция - номер заказа, то есть в таблице их может быть несколько.

Задания:**1. Нормализация данных:**

- Определите таблицы, которые необходимо создать для хранения информации о товарах, поставщиках и заказах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения всех заказов с указанием количества уникальных товаров в заказе.

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 5

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_5

Задача 1.

В магазине 20% товаров - электроника, 50% - одежда, 30% - продукты.
4% электроники, 12% одежды и 3% продуктов имеют дефекты. Какова вероятность выбрать дефектный товар?

В ответ запишите развернутое решение. Для обозначения вероятности воспользуйтесь $P(A)$ и $P(A|B)$.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

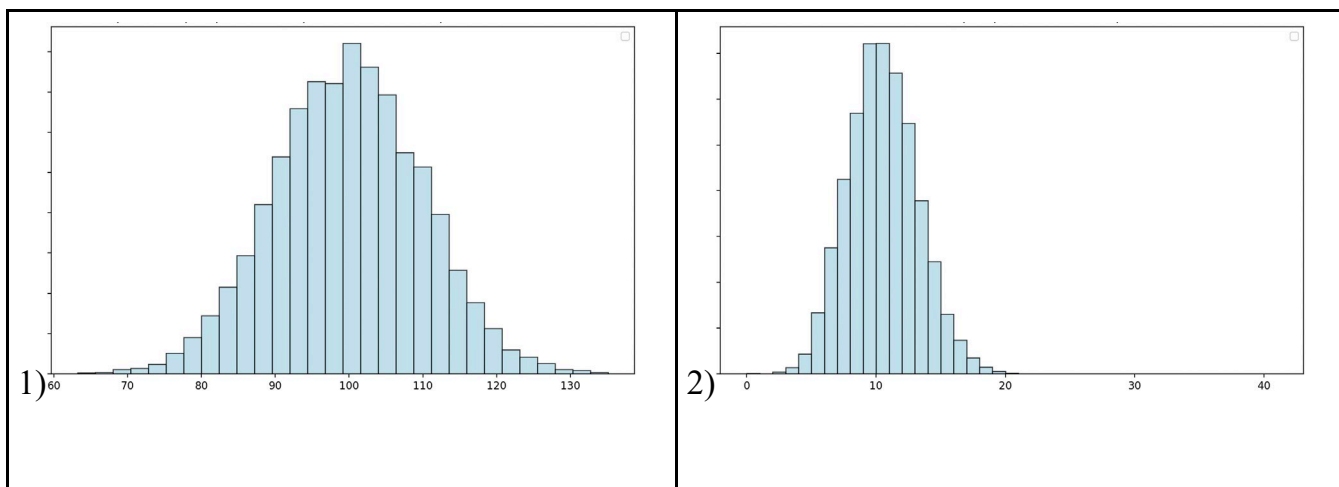
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

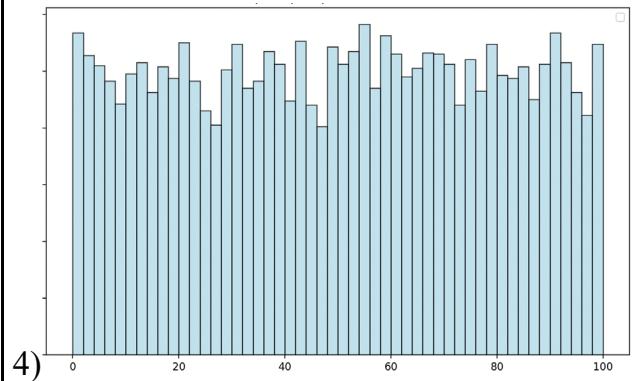
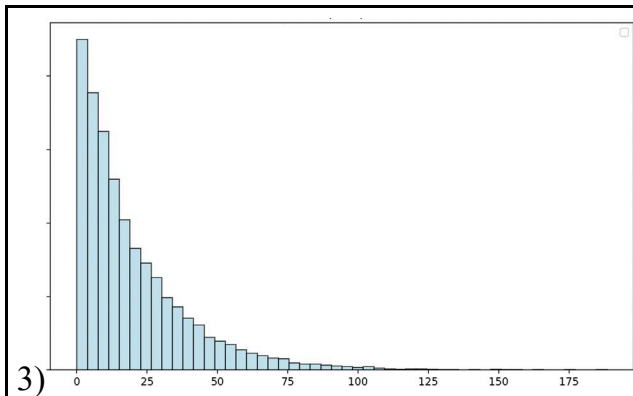
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Распределение доходов в большом городе: В крупном мегаполисе распределение доходов среди жителей разнится, при этом большинство людей зарабатывают средний доход, а меньшинство — значительно меньше или больше.

Б) Распределение числа студентов, записавшихся на онлайн-курс из группы в 40 человек, при вероятности регистрации 0,25.

В) Распределение времени до следующего сбоя в коммуникационной сети: Среднее время, которое проходит между двумя сбоями в коммуникационной сети равно 20 минут.

Г) Распределение характеризует уровень воды в резервуаре 100 мл.: Уровень воды варьируется равномерно между минимальным и максимальным значением.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая заменяет все цифры в строке на символ #.

Формат вывода: Строка после замены: <новая строка>.

Пример:

Ввод:

test123

Вывод:

Строка после замены: test###.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма, position - должность, age - возраст сотрудника.

Задача: Найти средний возраст сотрудников для каждого отдела.
Формат вывода: Название отдела, средний возраст работников

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (train_station.csv), содержащий информацию о железнодорожных рейсах. Этот набор данных включает следующие столбцы:

1. Номер поезда
2. Железнодорожная компания
3. Пункт назначения
4. Дата отправления
5. Время отправления
6. Дата прибытия
7. Время прибытия
8. Количество пассажиров
9. Задержка (в минутах)

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Найдите поезда, которые задерживались более чем на 30 минут.
2. Постройте столбчатую диаграмму, отображающую количество таких поездов по пунктам назначения.

Формат вывода:

Количество поездов с задержкой более 30 минут: <X>

Решение задачи и полученные файлы сохраните в папке **задача_5**.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код обучает линейную регрессию, но неправильно выводит оценку точности. Найдите и исправьте ошибку.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LinearRegression

X = [[1], [2], [3], [4]]
y = [2, 4, 6, 8]

model = LinearRegression()
model.fit(X, y)
print("Score:", model.score(X))
```

Код (решение) запишете в файл task_6.py или задача_6.txt

Задача 7.

Вы проектируете базу данных для системы управления заказами в ресторане.

Условия задачи:

- **Блюда:** Каждое блюдо имеет уникальный идентификатор, название, цену и категорию (супы, салаты и т.д.).

- **Заказы:** Каждый заказ имеет уникальный идентификатор, дату заказа, номер транзакции и блюдо, официанта.
- **Официанты:** Каждый официант имеет уникальный идентификатор, имя, фамилию и номер телефона.

Транзакции - номер заказа, то есть в таблице их может быть несколько.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о блюдах, заказах и официантах. (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Написать SQL-запрос, который вернет список официантов, дату и количество заказов, обслуженных каждым официантом на конкретные даты. Данные должны быть сгруппированы по официантам и датам.
-

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 6

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_6

Задача 1.

В лотерее разыгрываются 3 приза из 10 билетов. Какова вероятность выиграть хотя бы один приз, если у вас куплено 2 билета?

В ответ запишите развернутое решение. Для обозначения вероятности воспользуйтесь $P(A)$, а для обозначения сочетаний $C(n, k)$.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

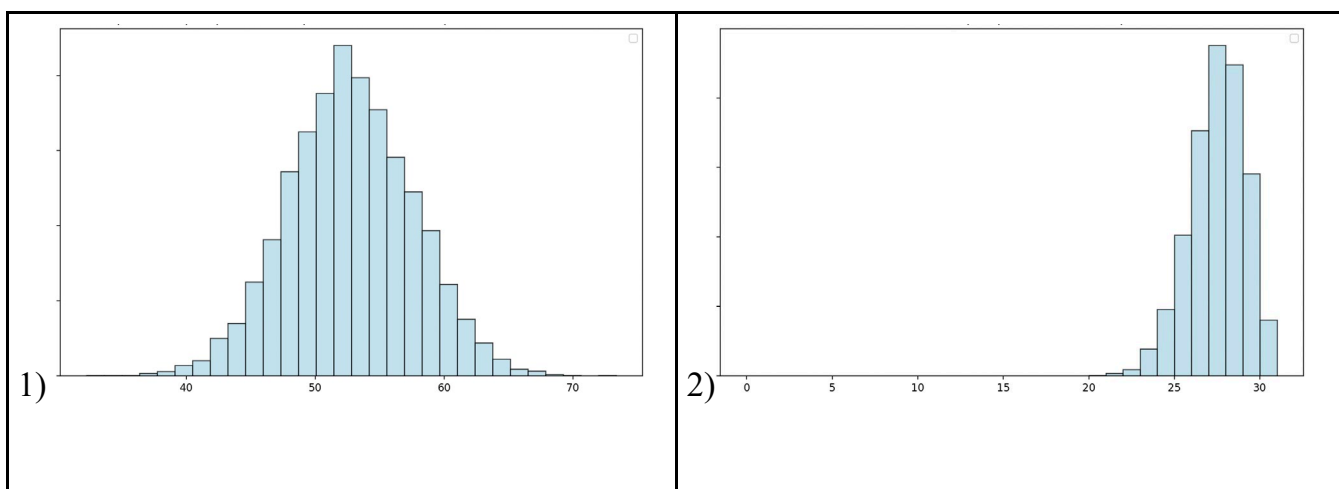
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

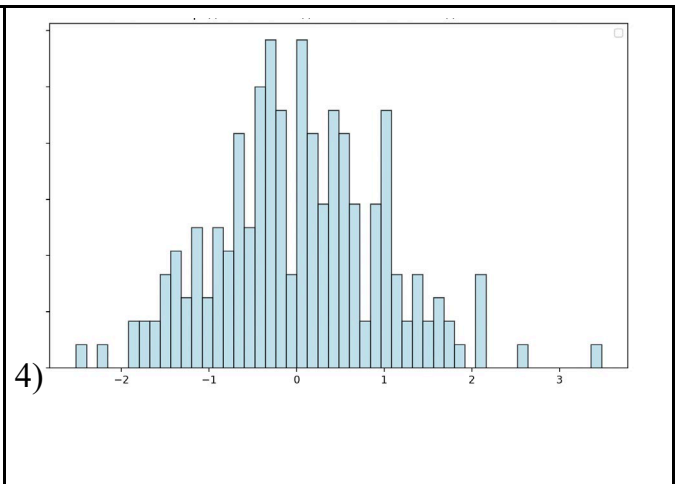
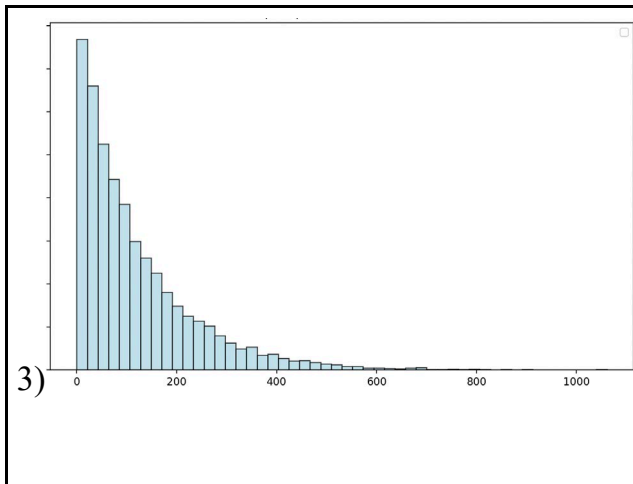
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Оценка среднего уровня токсичности нового препарата в небольшой группе пациентов: Применяется для анализа средних значений токсичности препарата при малом числе наблюдений

Б) Распределение времени до поломки машины: Среднее время, которое проходит до следующей поломки машины после текущего ремонта, равно 120 минут.

В) Распределение числа дней без аварий на производстве из 30 дней, при вероятности успешного дня 0,9.

Г) Распределение времени на выполнение теста: В классе из 100 студентов время, затраченное на выполнение экзамена, может различаться. Большинство студентов завершат тест за среднее время, в то время как некоторые завершат его быстрее или медленнее.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая проверяет, содержится ли в строке подстрока abc.

Формат вывода: Подстрока abc найдена в строке <введенная строка>: <Да/Нет>.

Пример:

Ввод:

alphabet

Вывод:

Подстрока abc найдена в строке alphabet: Нет.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица products (id, name, price, category, stock_quantity, created_at), где name -наименование продукта, price- цена, category - категория продукта, stock_quantity - количество товара на складе, created_at - дата добавления продукта в систему (ГГГГ-ММ-ДД).

Задача: Вывести названия всех продуктов, у которых количество на складе меньше 10.

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (train_station.csv), содержащий информацию о железнодорожных рейсах. Этот набор данных включает следующие столбцы:

1. Номер поезда
2. Железнодорожная компания
3. Пункт назначения
4. Дата отправления
5. Время отправления
6. Дата прибытия
7. Время прибытия
8. Количество пассажиров
9. Задержка (в минутах)

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите пункт назначения, куда чаще всего отправлялись поезда за последний год.
2. Постройте линейный график, показывающий изменение количества пассажиров в в этот пункт назначения по месяцам.

Формат вывода:

Пункт назначения с наибольшим количеством поездов: <Название пункта> - количество поездов: <X>

Решение задачи и полученные файлы сохраните в папке задача_5 .

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код создает модель логистической регрессии, но выдает ошибку при обучении. Найдите и исправьте ошибку.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LogisticRegression
import numpy as np

X = np.array([[1, 2], [2, 3], [3, 4]])
y = np.array([0, 1, 0])

model = LogisticRegression()
model.fit(X, y.reshape(-1, 1))
```

Код (решение) запишите в файл `task_6.py` или `задача_6.txt`

Задача 7.

Вы проектируете базу данных для управления расписанием занятий в учебном заведении.

Условия задачи:

- **Предметы:** Каждый предмет имеет уникальный идентификатор, название и количество часов.

- **Преподаватели:** Каждый преподаватель имеет уникальный идентификатор, имя, фамилию, квалификацию, предмет, который преподает.
- **Группы:** Каждая группа имеет уникальный идентификатор и список предметов, которые она изучает.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о предметах, преподавателях и группах напишите (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка групп и количества всех предметов.

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 7

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_7

Задача 1.

В лотерее разыгрываются 3 приза из 14 билетов. Какова вероятность выиграть хотя бы один приз, если у вас куплено 3 билета?

В ответ запишите развернутое решение. Для обозначения вероятности воспользуйтесь $P(A)$, а для обозначения сочетаний $C(n, k)$.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

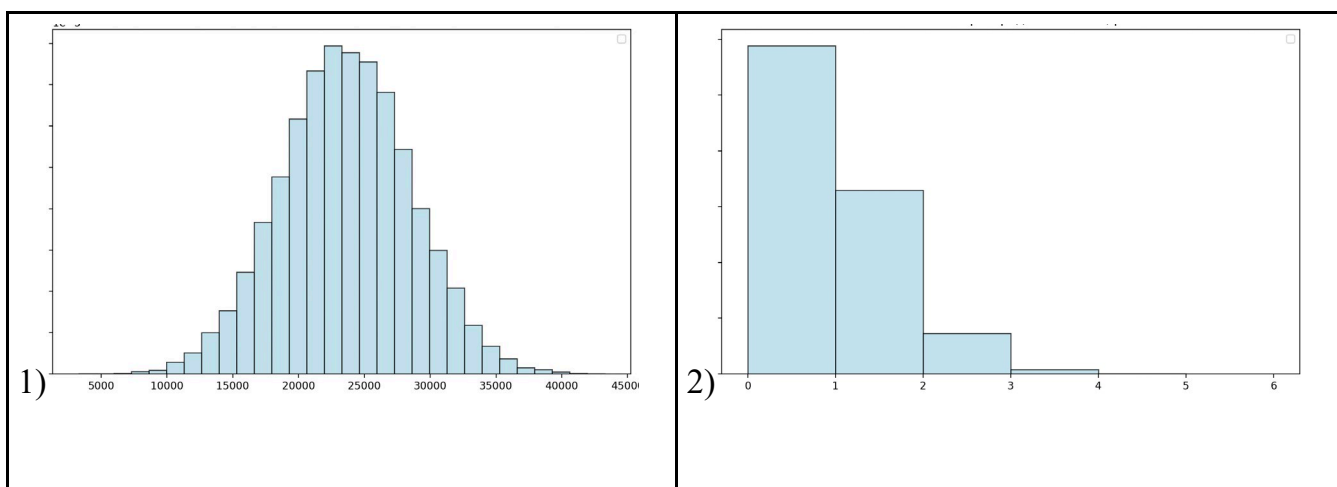
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

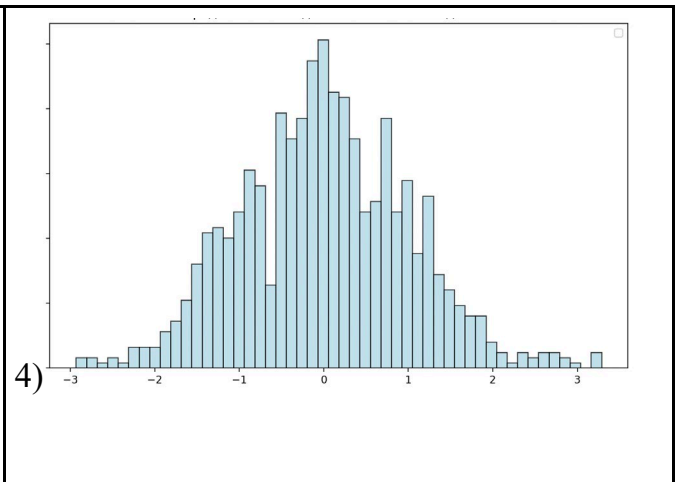
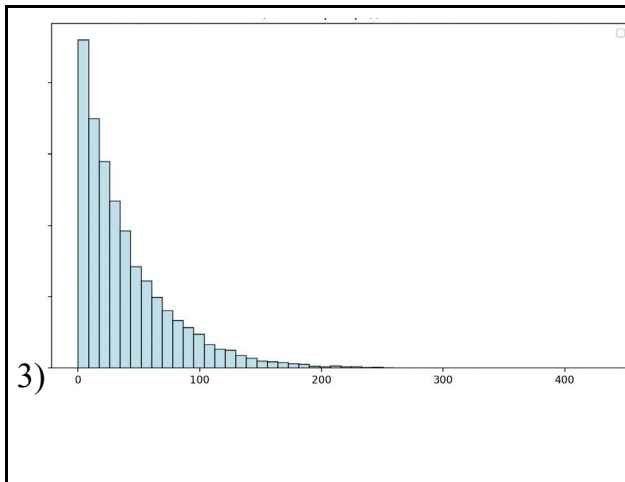
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Распределение времени до получения следующего вызова на мобильный телефон: Среднее время, которое проходит между двумя входящими вызовами на мобильный телефон, равно 40 минут.

Б) Оценка средних значений количества пользователей в небольших рекламных акциях: Используется для анализа среднего количества пользователей в различных акциях с малым числом наблюдений.

В) Распределение количества воды в озере в течение года: Если отслеживать объем воды в озере каждый день на протяжении года, то значения будут различаться, с наибольшим количеством наблюдений в среднем диапазоне и меньшим количеством — в крайних точках.

Г) Распределение числа выигрышных лотерейных билетов среди 5 купленных, при вероятности выигрыша 0,1.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая проверяет, является ли введенная строка палиндромом.

Формат вывода: Строка <введенная строка> является палиндромом: <Да/Нет>.

Пример:

Ввод:

radar

Вывод:

Строка radar является палиндромом: Да.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица products (id, name, price, category, stock_quantity, created_at), где name -наименование продукта, price- цена, category - категория продукта, stock_quantity - количество товара на складе, created_at - дата добавления продукта в систему (ГГГГ-ММ-ДД).

Задача: Выбрать общее количество товаров в каждой категории.

Формат вывода: Категория, количество товаров

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (library_books.csv), содержащий информацию о книгах в библиотеке. Этот набор данных включает следующие столбцы:

- * Название книги
- * Автор
- * Год издания
- * Жанр
- * Количество выданных за последний год
- * Общее количество экземпляров

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите автора, чьи книги были выданы чаще всего за последний год.
2. Постройте круговую диаграмму, показывающую распределение количества выданных книг для пяти самых популярных авторов. Убедитесь, что график имеет название и подписанные оси.

Формат вывода:

Автор, чьи книги были выданы чаще всего: "Имя автора" - количество выданных экземпляров: X

Решение задачи и полученные файлы сохраните в папке **задача_5**.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код пытается обучить логистическую регрессию, но выдает ошибку при вычислении метрик. Исправьте ошибку.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LogisticRegression

X = [[1, 2], [3, 4], [5, 6]]
y = [0, 1, 0]

model = LogisticRegression()
model.fit(X, y)
print("Accuracy:", model.score(X, y.reshape(-1, 1)))
```

Код (решение) запишете в файл `task_6.py` или `задача_6.txt`

Задача 7.

Вы проектируете базу данных для управления бронированием билетов в авиакомпании.

Условия задачи:

- **Рейсы:** Каждый рейс имеет уникальный идентификатор, маршрут, дату и время вылета, модель самолета.
- **Самолеты:** Каждый самолет имеет уникальный идентификатор, модель и количество мест.
- **Билеты:** Каждый билет имеет уникальный идентификатор, дату бронирования и ссылку на рейс и пассажира.
- **Пассажир:** Каждый пассажир имеет уникальный идентификатор, имя, фамилия, паспортные данные.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о рейсах, самолетах и билетах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка всех рейсов с указанием самолётов и пассажиров.

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 8

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_8

Задача 1.

В группе из 16 человек нужно выбрать 7 для участия в соревнованиях. Сколько различных способов существует для выбора команды?

В ответ запишите развернутое решение. Для обозначения сочетаний воспользуйтесь следующей записью - $C(n, k)$.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

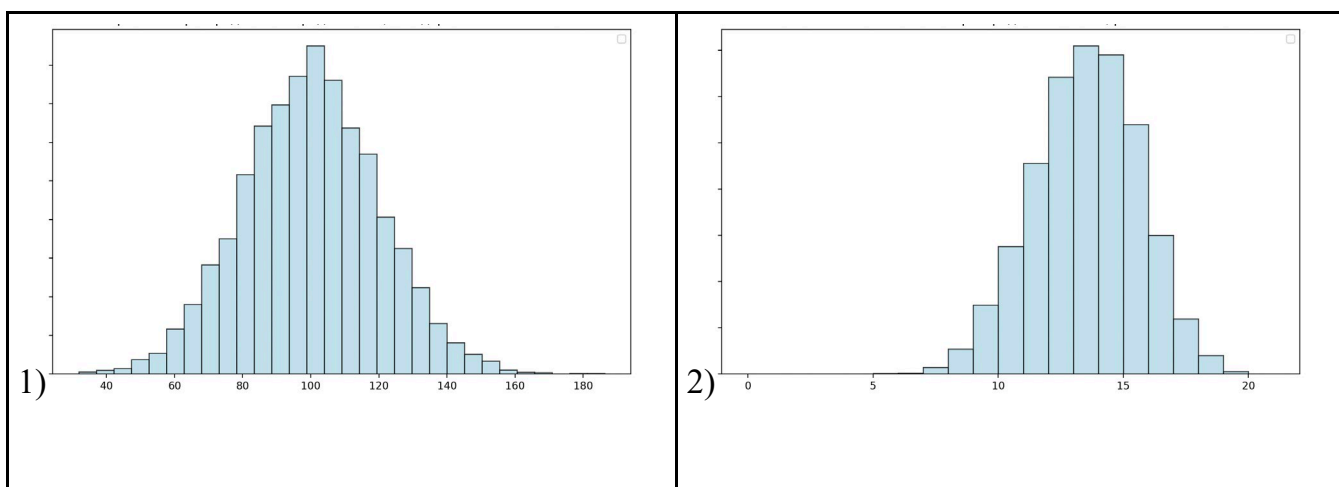
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

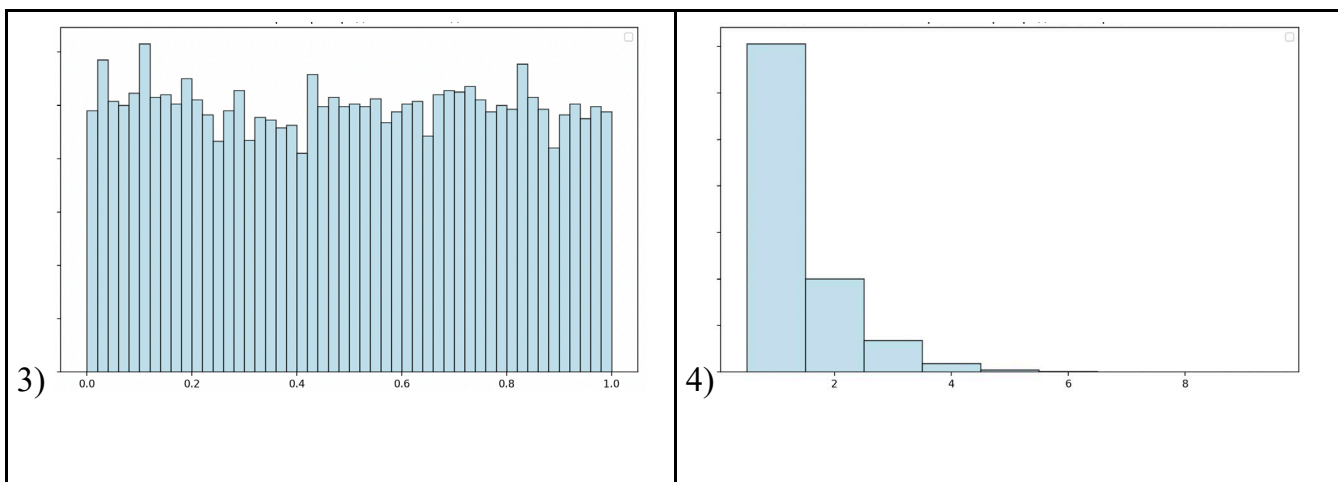
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Случайное распределение чисел от 0 до 1: Числа равномерно распределены в интервале от 0 до 1.

Б) Распределение числа голосов за кандидата в группе из 20 избирателей, где вероятность того, что избиратель поддержит кандидата, составляет 0,65.

В) Распределение числа клиентов в кафе за день: Если изучать количество клиентов, посещающих кафе ежедневно в течение месяца, их количество может колебаться, с большинством дней, когда приходит среднее количество посетителей

Г) Распределение описывает количество тестов до первой успешной проверки в системе качества: Число тестов или проверок, необходимых до первой успешной проверки в системе контроля качества.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая находит количество уникальных букв и символов в строке, введенной пользователем (игнорируйте регистр)

Формат вывода: f"Количество уникальных символов: {count_unique}"

Запишите свое решение в файл: *task_3.py*

Задача 4.

Вам дана таблица `products` (`id`, `name`, `price`, `category`, `stock_quantity`, `created_at`), где `name` -наименование продукта, `price`- цена, `category` - категория продукта, `stock_quantity` - количество товара на складе, `created_at` - дата добавления продукта в систему (ГГГГ-ММ-ДД).

Задача: Найти продукты, у которых количество на складе равно 0 (закончились).

Формат вывода: Наименование продукта

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (`library_books.csv`), содержащий информацию о книгах в библиотеке. Этот набор данных включает следующие столбцы:

- * Название книги
- * Автор
- * Год издания
- * Жанр
- * Количество выданных за последний год
- * Общее количество экземпляров

Используя библиотеку `pandas`, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Найдите книги, которые никогда не были выданы за последний год.
2. Постройте столбчатую диаграмму, отображающую количество таких книг по каждому жанру.

* Убедитесь, что график имеет название и подписанные оси.

Формат вывода:

Количество книг, которые не были выданы за последний год: X

Решение задачи и полученные файлы сохраните в папке **задача_5**.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
--------------	----------------	-----------------	----------------

1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Программа использует Pandas для вычисления среднего значения по столбцу, но выдает ошибку. Найдите и исправьте ошибку.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
import pandas as pd

data = {'A': [1, 2, 3, 4]}
df = pd.DataFrame(data)

mean_value = df.mean('A')
```

Код (решение) запишете в файл task_6.py или задача_6.txt

Задача 7.

Вы проектируете базу данных для платформы онлайн-обучения.

Условия задачи:

- **Курсы:** Каждый курс имеет уникальный идентификатор, название, дату начала и стоимость. Каждый курс может вести один или несколько преподавателей.

- **Преподаватели:** Каждый преподаватель имеет уникальный идентификатор, имя, фамилию и квалификацию.

• **Студенты:** Каждый студент имеет уникальный идентификатор, имя, фамилию, и дату регистрации. Один студент может записаться на несколько курсов.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о курсах, преподавателях и студентах. (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка всех курсов с указанием количества их преподавателей.

Поля для вывод: Курс, Кол-во

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 9

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_9

Задача 1.

В группе из 9 человек нужно выбрать 3 для участия в соревнованиях. Сколько различных способов существует для выбора команды?

В ответ запишите развернутое решение. Для обоснования воспользуйтесь символами: $C(n, k)$ - для обозначения сочетаний.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

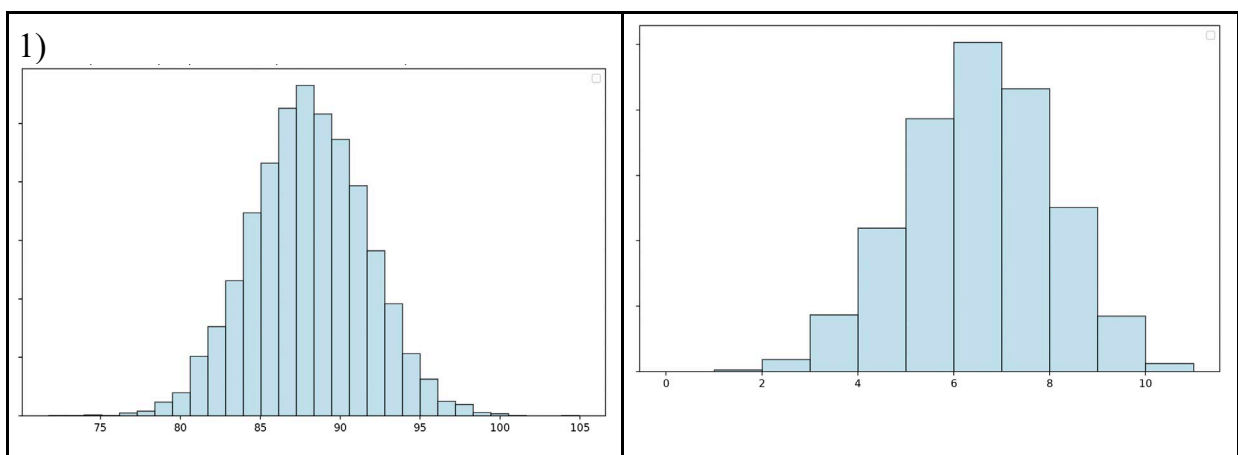
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

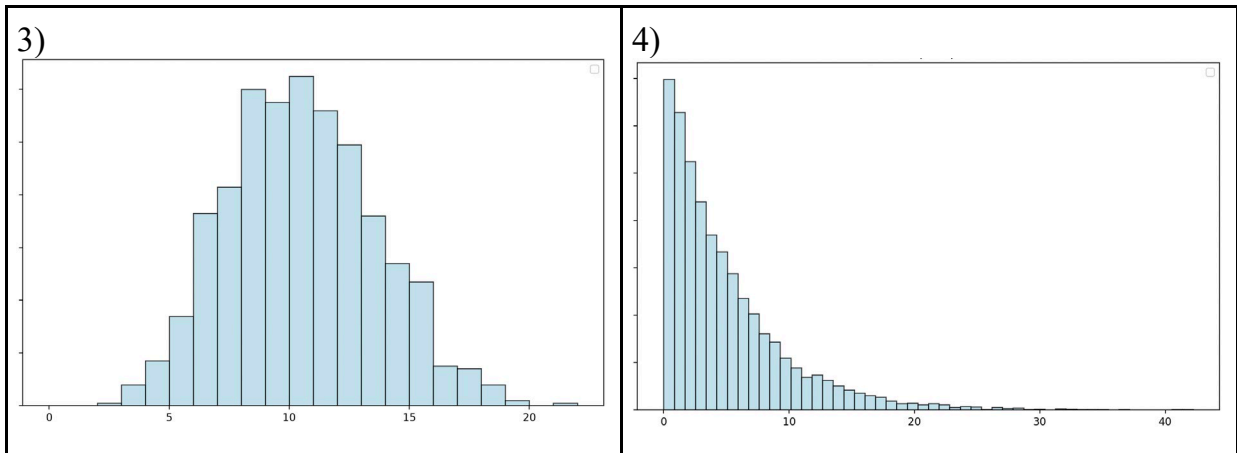
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Число звонков в службу поддержки за час: Среднее количество звонков в службу поддержки составляет 10 звонков в час.

Б) В большой выборке населения уровни IQ варьируются, при этом большинство людей имеют средний уровень интеллекта, а меньшинство — более низкий или высокий уровень.

В) Распределение времени ожидания автобуса: Среднее время, которое проходит между прибытиями автобусов на определённой остановке, равно 5 минут.

Г) Распределение числа успешных бросков мяча в корзину из 10 попыток, где вероятность успешного броска составляет 0,6.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая вычисляет сумму элементов списка целых чисел.

Формат ввода:

Целые числа, разделенные звездочкой.

Формат вывода: Сумма элементов списка: <сумма>

Пример:

Ввод:

1*2*3*4*5

Вывод:

Сумма элементов списка: 15

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма (ГГГГ-ММ-ДД), position - должность, age - возраст сотрудника.

Задача: найти сотрудников, чья зарплата составляет более 50000.

Формат вывода: Имена сотрудников, дата найма

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (sports_competition.csv), содержащий информацию о спортивных соревнованиях. Этот набор данных включает следующие столбцы:

1. Название соревнования
2. Участник
3. Страна
4. Год проведения
5. Вид спорта
6. Количество набранных очков
7. Количество медалей

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите топ-4 спортсменов, набравших наибольшее количество очков за три года с 2019 года до 2021 года.
2. Постройте круговую диаграмму, показывающую распределение медалей среди этих четырех спортсменов.

Формат вывода:

Топ-4 спортсменов по количеству набранных очков:

<Имя спортсмена 1> (<Страна>) - количество очков: X

<Имя спортсмена 2> (<Страна>) - количество очков: Y

<Имя спортсмена 3> (<Страна>) - количество очков: Z

<Имя спортсмена 4> (<Страна>) - количество очков: Т

Решение задачи и полученные файлы сохраните в папке **задача_5**.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	Диаграмма, полученная в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Вам дан код для линейной регрессии. В данном коде есть ошибка - найдите ее и исправьте.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LinearRegression
import numpy as np

# Генерация неправильных данных
X = np.array([[1], [2], [3], [4]])
y = np.array([1.5, 1.7, 1.9, 2.1, 3.4])

# Обучение модели
model = LinearRegression()
model.fit(X, y)

# Предсказания
predictions = model.predict(X)
print(predictions)
```

Код (решение) запишете в файл `task_6.py` или `задача_6.txt`

Задача 7.

Вы работаете в компании, управляющей сетью кинотеатров. Необходимо спроектировать базу данных для хранения информации о фильмах, сеансах и билетах.

Условия задачи:

- * Фильмы: Каждый фильм имеет уникальный идентификатор, название, жанр и продолжительность.
- * Сеансы: Каждый сеанс привязан к одному фильму, имеет уникальный идентификатор, дату и время показа, зал.
- * Билеты: Каждый билет имеет уникальный идентификатор, привязан к сеансу и месту в зале, включает цену.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о фильмах, сеансах и билетах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения информации о всех проданных билетах для 4-го сеанса, включая данные о месте, цене билета, времени сеанса и зале.

Требуемые поля

- Время сеанса (Screening Time),
- Зал (Hall),
- Место (Seat),
- Цена билета (Price).

Примеры возможного заполнения таблиц:

Таблица Фильмы:

id	title	genre	duration
1	Inception	Sci-Fi	148
2	Titanic	Drama	195
3	The Matrix	Action	136
4	Avatar	Sci-Fi	162
5	The Godfather	Crime	175

Таблица Сеансы:

id	movie_id	screening_time	hall
1	1	2024-09-10 18:00:00	Hall 1
2	2	2024-09-10 20:00:00	Hall 2
3	3	2024-09-11 15:30:00	Hall 3
4	4	2024-09-11 19:00:00	Hall 1
5	5	2024-09-12 21:00:00	Hall 2

Таблица Билетов:

id	screening_id	seat	price
1	1	A1	10.50
2	1	A2	10.50
3	2	B3	12.00
4	2	B4	12.00
5	3	C5	11.00
6	4	D1	13.50
7	5	E2	14.00

Решение задачи и полученные файлы сохраните в папке задача_7
Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 10

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_10

Задача 1.

В компании 70% сотрудников - программисты, а 30% - дизайнеры. 80% программистов имеют высшее образование, а 50% дизайнеров - нет. Какова вероятность того, что случайно выбранный сотрудник имеет высшее образование?

В ответ запишите развернутое решение. Для обоснования воспользуйтесь символами: $P(A)$ - для обозначения вероятности.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

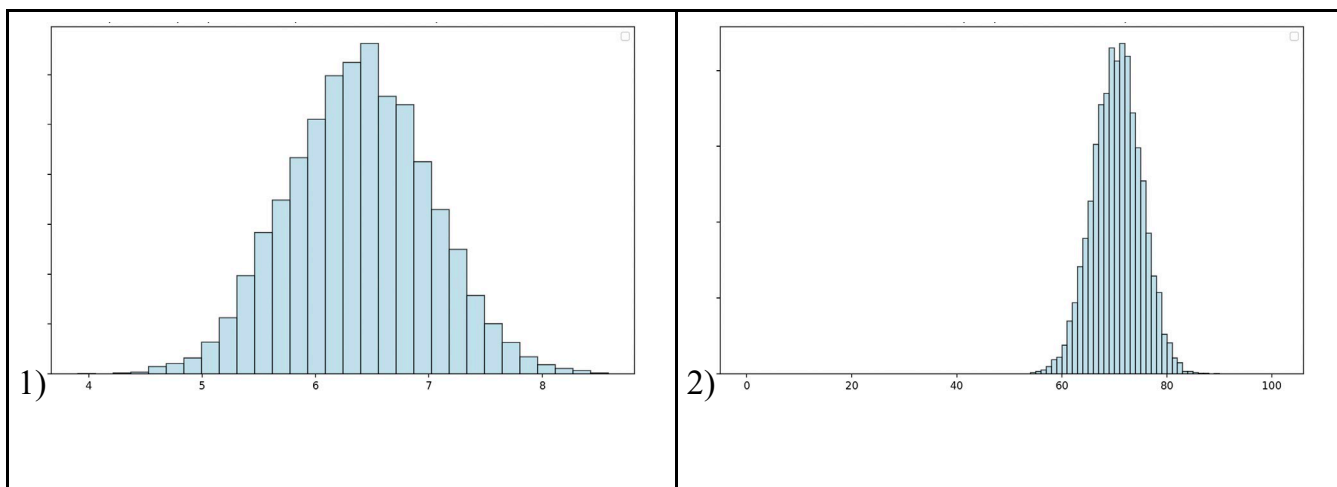
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

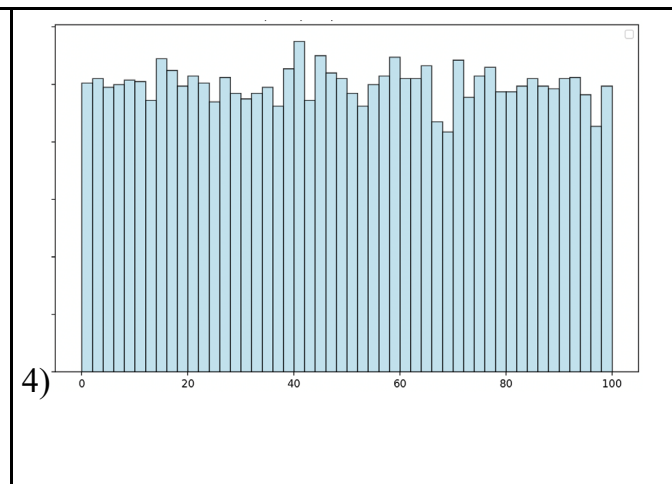
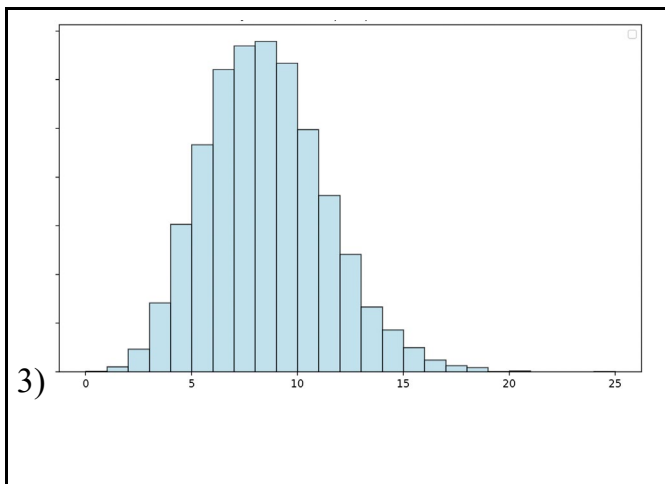
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Распределение скорости ветра в течение года на одной метеостанции: Ежегодное распределение скорости ветра на определенной метеостанции, будет варьироваться, с преобладанием средних значений и редкими крайними отклонениями.

Б) Число сбоев в работе серверов за месяц: Среднее количество сбоев в работе серверов за месяц составляет 8.

В) Распределение чисел на рулетке. Числа на колесе рулетки распределены равномерно, при этом каждый номер имеет равную вероятность выпадения.

Г) Распределение числа пациентов с положительной реакцией на лекарство из 100 испытуемых, при вероятности успеха 0,7.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая вычисляет сумму всех чисел от 1 до N (включительно), где N (целое число) вводится с клавиатуры.

Формат вывода: Сумма чисел от 1 до <N>: <сумма>.

Пример:

Ввод:

5

Вывод:

Сумма чисел от 1 до 5: 15.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма, position - должность, age - возраст сотрудника.

Задача: Отсортировать сотрудников по зарплате в порядке убывания.

Формат вывода: Имя сотрудника, зарплата.

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (sports_competition.csv), содержащий информацию о спортивных соревнованиях. Этот набор данных включает следующие столбцы:

1. Название соревнования
2. Участник
3. Страна
4. Год проведения
5. Вид спорта
6. Количество набранных очков
7. Количество медалей

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите вид спорта, в котором спортсмены набрали наибольшее количество очков за все время.
2. Постройте линейный график, показывающий изменение количества очков по годам проведения соревнований в этом виде спорта.

Формат вывода:

Вид спорта с наибольшим количеством набранных очков: <Название вида спорта>
- количество очков: <X>

Решение задачи и полученные файлы сохраните в папке **задача_5**.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код ниже можно использовать для реализации метода К-ближайших соседей.

Найдите ошибку в коде и исправьте её

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier

# Загружаем набор данных
iris = load_iris()
X, y = iris.data, iris.target

# Разделяем данные
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

# Создание модели KNN
knn = KNeighborsClassifier(n_neighbors=-3)
knn.fit(X_train, y_train)

# Оценка точности
accuracy = knn.score(X_test, y_test)
print("Точность:", accuracy)
```

Код (решение) запишете в файл task_6.py или задача_6.txt

Задача 7.

Вы разрабатываете базу данных для автосалона, который продает автомобили и ведет учет клиентов и продаж.

Условия задачи:

- * Машины: Автомобили имеют уникальный идентификатор, модель, производителя и цену.
- * Клиенты: Каждый клиент имеет уникальный идентификатор, имя и контактную информацию.
- * Продажа: Продажи автомобилей ведутся с указанием даты продажи и клиента, который купил автомобиль.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации автомобилях, клиентах и продажах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка всех продаж, включающего следующие данные:
 - Идентификатор продажи
 - Дата продажи
 - Модель автомобиля
 - Имя клиента
 - Контактные данные клиента

Используйте таблицы Sales, Cars и Clients, которые связаны между собой через идентификаторы автомобиля и клиента.

Решение задачи и полученные файлы сохраните в папке задача_7
Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 11

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_11

Задача 1.

Даны значения: 2, 8, 3, 10, 1. Какова медиана этих данных?

В ответ запишите развернутое решение.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

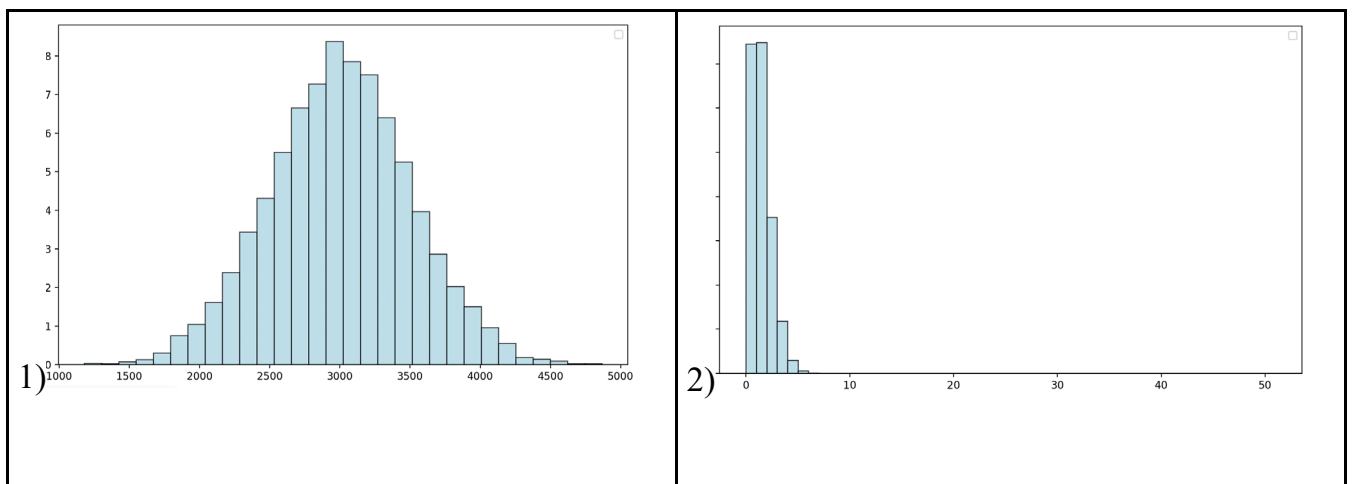
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

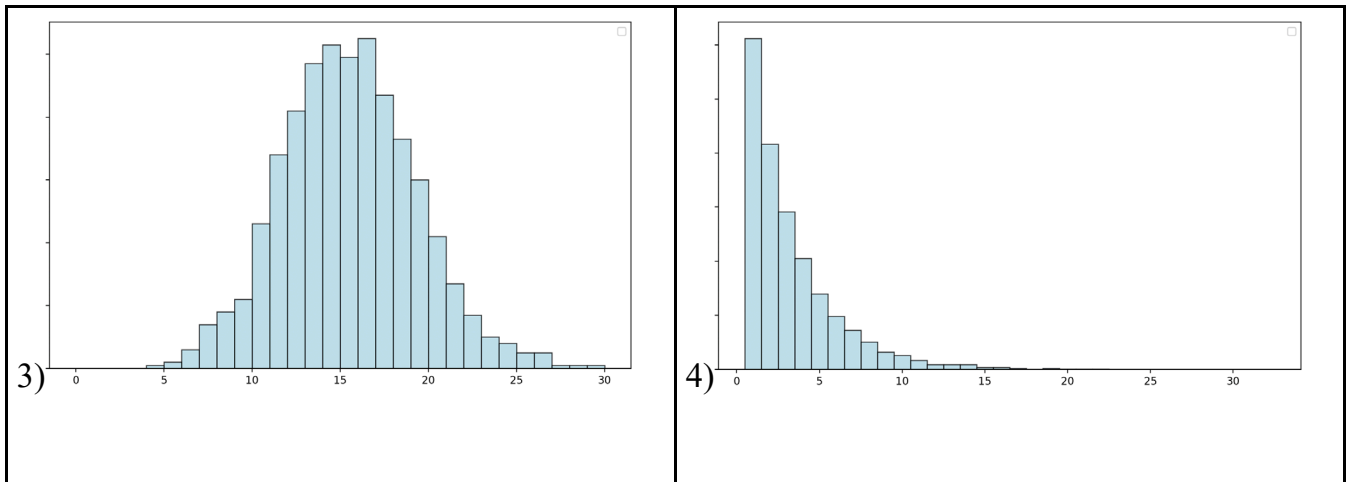
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Распределение числа дефектных изделий в партии из 50 штук, при вероятности дефекта 0,02.

Б) Число автомобилей, проезжающих через перекресток за минуту: Среднее количество автомобилей, проезжающих через перекресток, равно 15 машин в минуту.

В) Распределение описывает количество попыток при бросании монеты до первого орла: Число бросков монеты, необходимых, чтобы впервые выпал орел.

Г) Распределение веса новорожденных: В крупной больнице вес новорожденных детей варьируется, при этом большинство детей имеют средний вес, а меньшинство — существенно ниже или выше среднего.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая находит и выводит позицию первого вхождения подстроки `end` в строке. Если подстрока не найдена, выведите - `end`.

Формат вывода: Позиция первого вхождения подстроки `end` в строке <введенная строка>: <позиция>.

Пример:

Ввод:

abcdefend

Вывод:

Позиция первого вхождения подстроки end в строке abcdefend: 6.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма, position - должность, age - возраст сотрудника.

Задача: Найти минимальную зарплату среди всех сотрудников.

Вывод: Минимальная зарплата

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (airport_flights.csv), содержащий информацию рейсах в аэропорту. Этот набор данных включает следующие столбцы:

1. Номер рейса
2. Авиакомпания
3. Пункт назначения
4. Время вылета
5. Время прилета
6. Количество пассажиров
7. Задержка (в минутах)

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Найдите рейсы, которые задерживались менее чем на 1 час за последний месяц.
2. Постройте столбчатую диаграмму, отображающую количество таких рейсов по пунктам назначения.

Формат вывода:

Количество рейсов с задержкой менее 1 часа: <X>

Решение задачи и полученные файлы сохраните в папке задача_5.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
--------------	----------------	-----------------	----------------

1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

В следующем коде производится оценка модели, но в нем есть несколько ошибок. Найдите и исправьте их.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import mean_squared_error

# Загрузка данных
iris = load_iris()
X, y = iris.data, iris.target

# Разделяем данные
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

# Обучаем модель
model = DecisionTreeClassifier()
model.fit(X_train, y_train)

# Предсказания
y_pred = model.predict(X_test)

# Оценка качества
mse = mean_squared_error(y_test, y_pred)
print("Средняя ошибка:", mse)
```

Код (решение) запишете в файл task_6.py или задача_6.txt

Задача 7.

Вы создаете базу данных для ресторана, чтобы хранить информацию о блюдах, клиентах и заказах.

Условия задачи:

1. Блюда имеют уникальный идентификатор, название и цену.
2. Клиенты ресторана имеют уникальный идентификатор, имя и контактную информацию.
3. Заказы включают несколько блюд, и каждый заказ имеет уникальный идентификатор, дату, номер чека, имя клиента и общую сумму.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о блюдах, клиентах и заказах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка всех клиентов, которые заказали морковный торт.

Решение задачи и полученные файлы сохраните в папке задача_7
Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 12

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_12

Задача 1.

В коробке 9 красных, 4 зеленых и 5 синих шара. Если извлечь 2 шара, какова вероятность того, что оба шара будут красными?

В ответ запишите развернутое решение. Для обозначения числа сочетаний воспользуйтесь: $C(n, k)$, а для обозначения вероятности P .

Запишите свое решение в файл: задача_1.txt.

Задача 2.

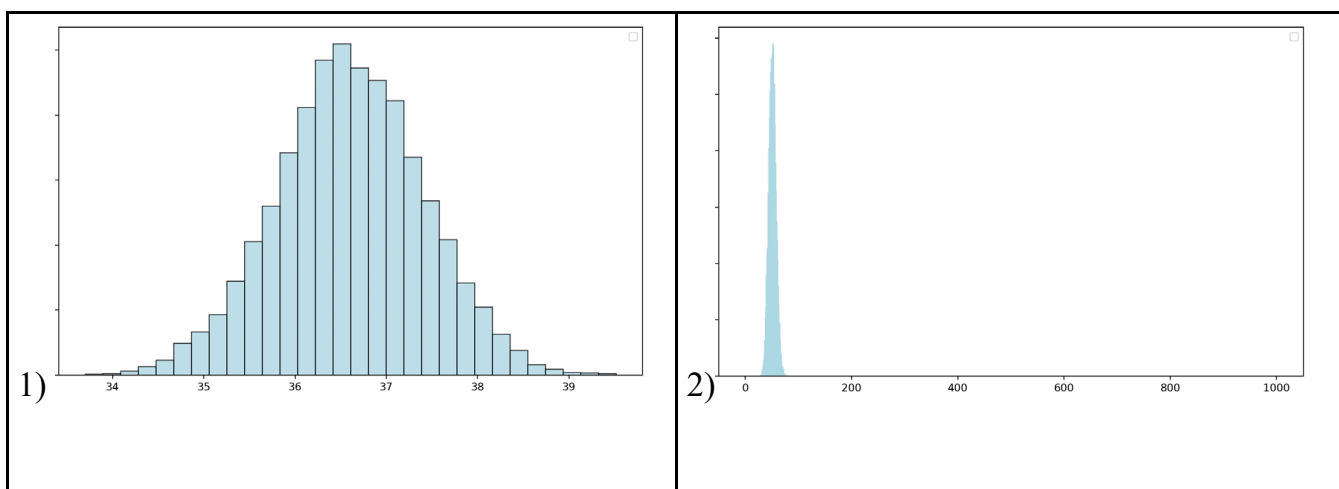
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

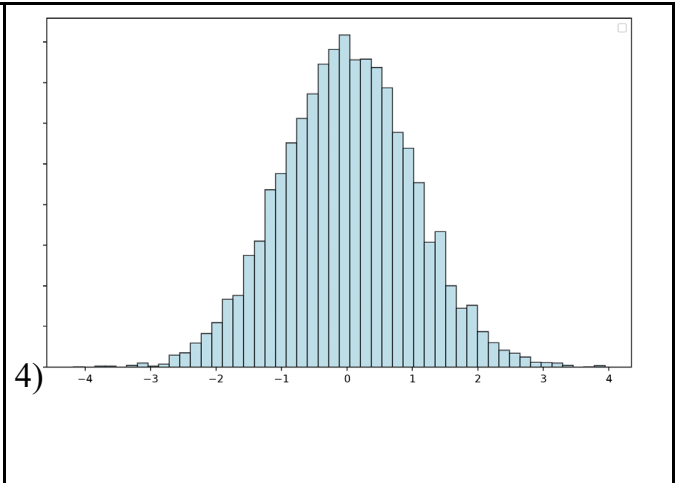
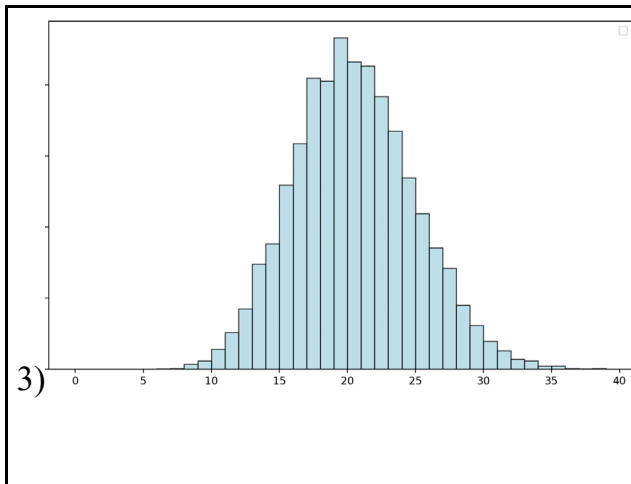
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Оценка среднего значения зарплаты сотрудников в небольшой компании: Используется при оценке среднего значения зарплаты, когда размер выборки небольшой и стандартное отклонение неизвестно.

Б) Число новых сообщений на электронной почте за день: Среднее количество новых сообщений на электронной почте за день равно 20 сообщений.

В) Распределение температуры тела у здоровых людей: Если измерить температуру тела у большой группы здоровых людей, она будет колебаться вокруг среднего значения, около $36,6^{\circ}\text{C}$, с небольшими отклонениями.

Г) Распределение числа пользователей, откликнувшихся на рекламную рассылку из 1000 отправленных писем, при вероятности отклика 0,05.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая заменяет все вхождения подстроки duck на Horse в введенной строке.

Формат вывода: Строка после замены: <новая строка>.

Пример:

Ввод:

duck and ducklog

Вывод:

Строка после замены: Horse and Horselog.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма, position - должность, age - возраст сотрудника.

Задача: Вывести имена всех сотрудников, которые работают на позиции «Manager» и их возраст составляет от 20 до 35 лет.

Вывод: имена сотрудников

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (airport_flights.csv), содержащий информацию рейсах в аэропорту. Этот набор данных включает следующие столбцы:

1. Номер рейса
2. Авиакомпания
3. Пункт назначения
4. Время вылета
5. Время прилета
6. Количество пассажиров
7. Задержка (в минутах)

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите топ-3 авиакомпаний по наименьшему количеству задержек.
2. Постройте круговую диаграмму, показывающую распределение количества пассажиров среди этих авиакомпаний.

Формат вывода:

Топ-3 авиакомпаний по количеству наименьшего количества задержек:

<Авиакомпания 1> - количество задержек: X

<Авиакомпания 2> - количество задержек: Y

<Авиакомпания 3> - количество задержек: Z

Решение задачи и полученные файлы сохраните в папке **задача_5**.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код создает и обучает модель линейной регрессии, но выводит ошибку при обучении. Найдите ошибку и исправьте ее.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LinearRegression
import numpy as np

X = np.array([[1, 2], [3, 4], [5, 6]])
y = np.array([3, 6, 9])

model = LinearRegression()
model.fit(X.reshape(-1, 1), y)
```

Код (решение) запишете в файл `task_6.py` или `задача_6.txt`

Задача 7.

Вы работаете в компании, управляющей складом. Требуется база данных для учёта товаров, поставщиков и заказов на склад.

Условия задачи:

Товары:

- Каждый товар имеет уникальный идентификатор, название, цену и количество на складе.
- Товары связаны с поставщиками через уникальный идентификатор поставщика.

Поставщики:

- Каждый поставщик имеет уникальный идентификатор, название компании и контактные данные (телефон и электронную почту).

Заказы:

- Заказ может включать несколько товаров, но каждый товар может быть заказан только один раз в каждом заказе.
- У каждого заказа есть уникальный номер чека.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о товарах, поставщиках и заказах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения всех поставщиков с указанием средней цены на их товары.

Поля:

Название поставщика

Средняя

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 13

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_13

Задача 1.

В магазине 15% товаров - электроника, 70% - одежда, 15% - продукты. 3% электроники, 11% одежды и 7.5% продуктов имеют дефекты. Какова вероятность выбрать дефектный товар?

В ответ запишите развернутое решение. Для обозначения вероятности воспользуйтесь $P(A)$ и $P(A|B)$.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

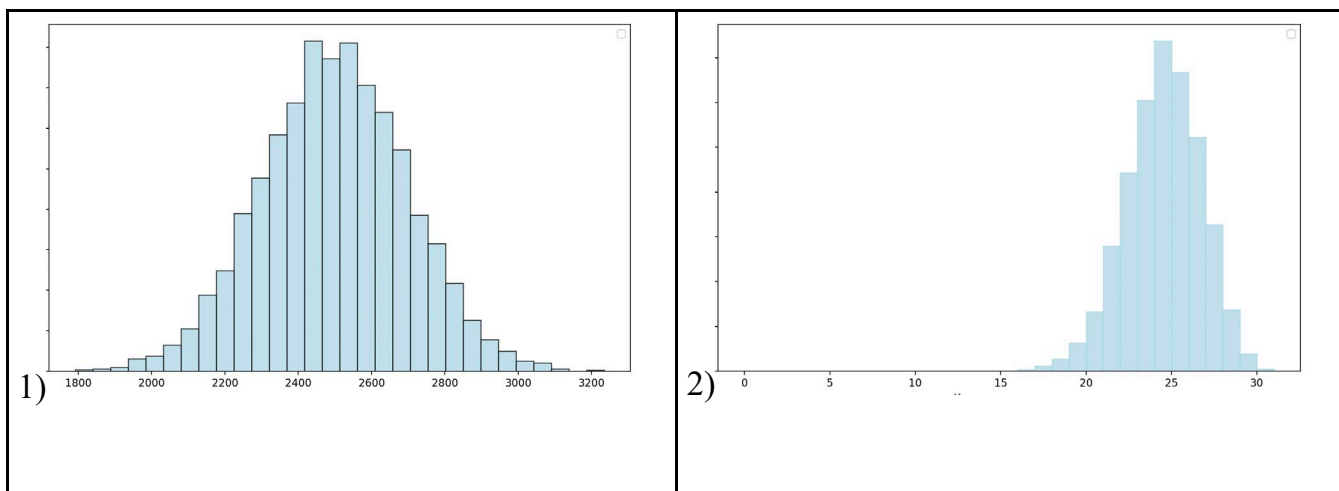
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

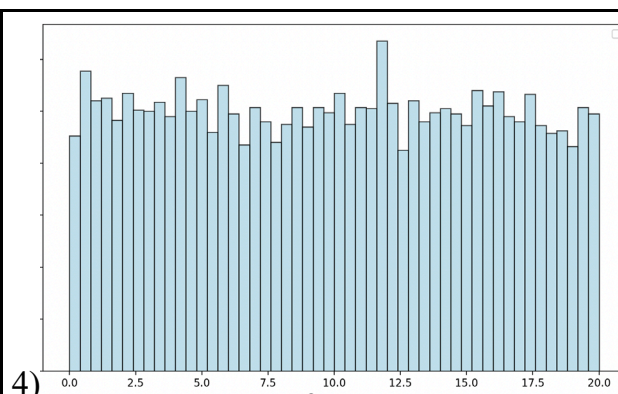
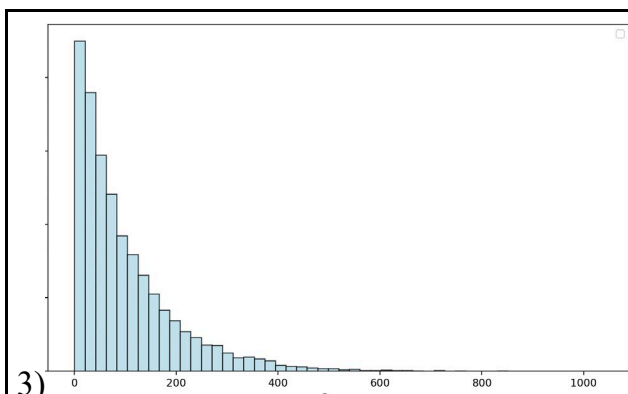
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Распределение высоты деревьев в лесу: В большом лесу высота деревьев одного и того же вида будет варьироваться, при этом большинство деревьев достигают среднего роста, а некоторые — заметно выше или ниже. (В распределении высота указана в см.)

Б) Распределение числа студентов, сдавших экзамен из 30 человек, при вероятности успеха 0,8.

В) Распределение времени жизни радиоактивного атома: среднее время, в течение которого радиоактивный атом остаётся нестабильным до его распада, равно 100 лет.

Г) Распределение выигрыша в простой лотерее: Выигрышные суммы в лотерее распределены равномерно между минимальной и максимальной возможной суммой.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая заменяет все цифры в строке на символ @.

Формат вывода: Строка после замены: <новая строка>.

Пример:

Ввод:

test123

Вывод:

Строка после замены: test@@@.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица employees (id, name, salary, department, hire_date, position, age), где name - имя сотрудника, salary - зарплата, hire_date - дата найма, position - должность, age - возраст сотрудника.

Задача: Найти минимальный возраст сотрудников для каждого отдела.
Формат вывода: Название отдела, минимальный возраст работников

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (train_station.csv), содержащий информацию о железнодорожных рейсах. Этот набор данных включает следующие столбцы:

1. Номер поезда
2. Железнодорожная компания
3. Пункт назначения
4. Дата отправления
5. Время отправления
6. Дата прибытия
7. Время прибытия
8. Количество пассажиров
9. Задержка (в минутах)

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Найдите поезда, которые задерживались менее чем на 2 минуты.
2. Постройте столбчатую диаграмму, отображающую количество таких поездов по железнодорожным компаниям.

Формат вывода:

Количество поездов с задержкой менее 2 минут: <X>

Решение задачи и полученные файлы сохраните в папке задача_5 .

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код обучает линейную регрессию, но неправильно выводит оценку точности. Найдите и исправьте ошибку.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LinearRegression

X = [[1], [2], [3], [4]]
y = [2, 4, 6, 8]

model = LinearRegression()
model.fit(X, y)
print("Score:", model.score(y, X))
```

Код (решение) запишете в файл `task_6.py` или `задача_6.txt`

Задача 7.

Вы проектируете базу данных для системы управления заказами в ресторане.

Условия задачи:

1. Блюда:

- У каждого блюда есть уникальный идентификатор, название и цена.
- Каждое блюдо принадлежит к категории (например, супы, салаты).

2. Заказы:

- Каждый заказ имеет уникальный идентификатор и дату.
- Заказ может содержать несколько блюд.
- Каждый заказ связан с официантом по его идентификатору.

3. Официанты:

- У каждого официанта есть уникальный идентификатор, имя и номер телефона.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о блюдах, заказах и официантах. (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения количества заказов у каждого официанта. Запрос должен возвращать имя официанта и количество его заказов.

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 14

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_14

Задача 1.

В лотерее разыгрываются 7 приза из 11 билетов. Какова вероятность выиграть хотя бы один приз, если у вас куплено 3 билета?

В ответ запишите развернутое решение. Для обозначения вероятности воспользуйтесь $P(A)$, а для обозначения сочетаний $C(n, k)$.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

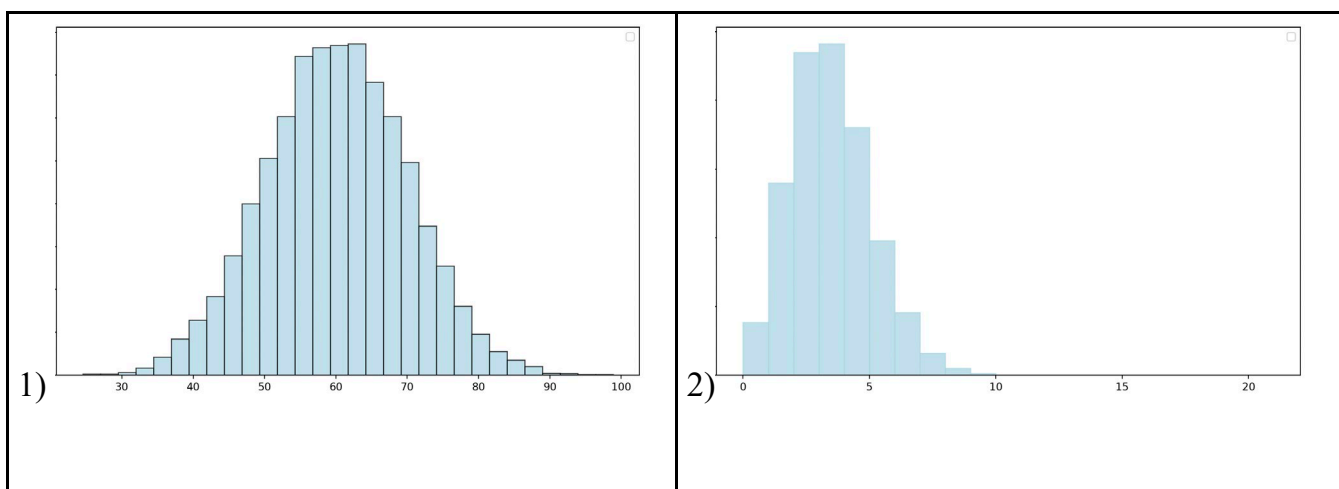
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

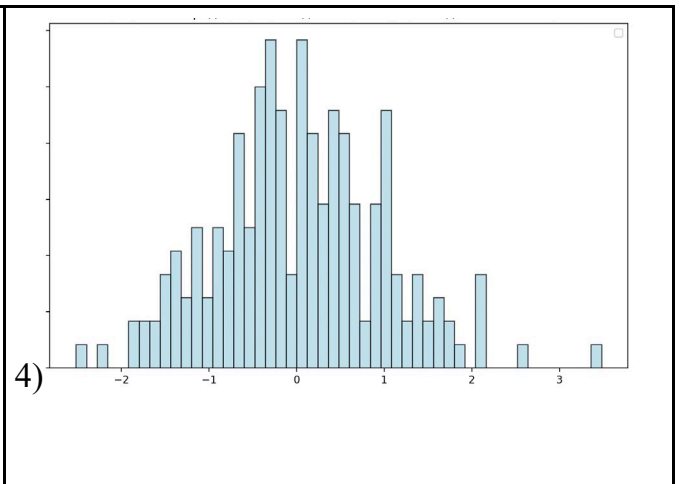
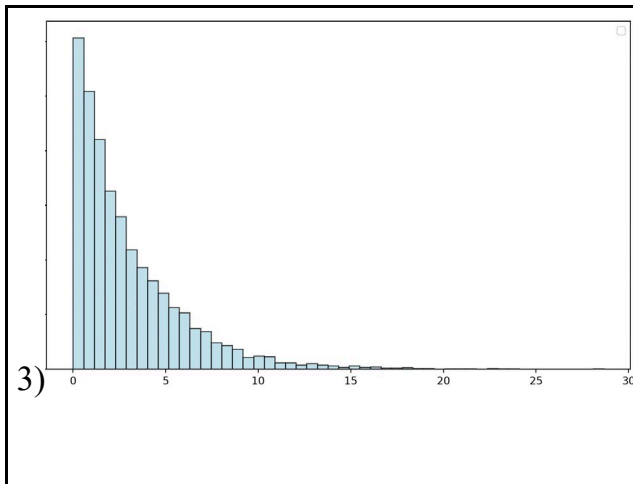
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Сравнение результатов двух небольших медицинских исследований: Используется для сравнения средних значений в двух группах с небольшим размером выборки.

Б) Распределение времени до следующего звонка в службу поддержки: Среднее время, которое проходит между поступлением двух звонков в службу поддержки, равное 3 минуты.

В) Распределение числа машин, нарушивших правила парковки на улице из 20 припаркованных, при вероятности нарушения 0,15.

Г) Распределение результатов ЕГЭ среди школьников: Если рассматривать результаты по одному предмету, то большинство школьников набирает среднее количество баллов, тогда как количество тех, кто набирает минимальные или максимальные баллы, гораздо меньше.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая проверяет, содержится ли в строке подстрока `dog`.

Формат вывода: Подстрока `dog` найдена в строке <введенная строка>: <Да/Нет>.

Пример:

Ввод:

alphabet

Вывод:

Подстрока dog найдена в строке alphabet: Нет.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица products (id, name, price, category, stock_quantity, created_at), где name -наименование продукта, price- цена, category - категория продукта, stock_quantity - количество товара на складе, created_at - дата добавления продукта в систему (ГГГГ-ММ-ДД).

Задача: Вывести названия всех продуктов, у которых цена больше 10000.

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (train_station.csv), содержащий информацию о железнодорожных рейсах. Этот набор данных включает следующие столбцы:

1. Номер поезда
2. Железнодорожная компания
3. Пункт назначения
4. Дата отправления
5. Время отправления
6. Дата прибытия
7. Время прибытия
8. Количество пассажиров
9. Задержка (в минутах)

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите пункт назначения, куда реже всего отправлялись поезда за последний год. Если таких несколько, то выведите первый по алфавиту.
2. Постройте круговую диаграмму, показывающую количество пассажиров прибывших поездов в пункты Rayfort, Crystalview, Hallbury.

Формат вывода:

Пункт назначения с наименьшим количеством поездов: <Название пункта> - количество поездов: <X>

Решение задачи и полученные файлы сохраните в папке задача_5 .

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
--------------	----------------	-----------------	----------------

1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код создает модель логистической регрессии, но выдает ошибку при обучении. Найдите и исправьте ошибку.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LogisticRegression
import numpy as np

X = np.array([[1, 2], [2, 3], [3, 4]])
y = np.array([0, 1, 0])

model = LogisticRegression()
model.fit(X.reshape(-1, 1), y)
```

Код (решение) запишете в файл task_6.py или задача_6.txt

Задача 7.

Вы проектируете базу данных для управления расписанием занятий в учебном заведении.

Условия задачи:

- **Предметы:** Каждый предмет имеет уникальный идентификатор, название и количество часов, преподавателя, который ведет этот предмет.

- **Преподаватели:** Каждый преподаватель имеет уникальный идентификатор, имя, фамилию и квалификацию.
- **Группы:** Каждая группа имеет уникальный идентификатор и список предметов, которые она изучает.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о предметах, преподавателях и группах напишите (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка групп и количества их предметов.

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 15

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_15

Задача 1.

В лотерее разыгрываются 5 приза из 23 билетов. Какова вероятность выиграть хотя бы один приз, если у вас куплено 3 билета?

В ответ запишите развернутое решение. Для обозначения вероятности воспользуйтесь $P(A)$, а для обозначения сочетаний $C(n, k)$.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

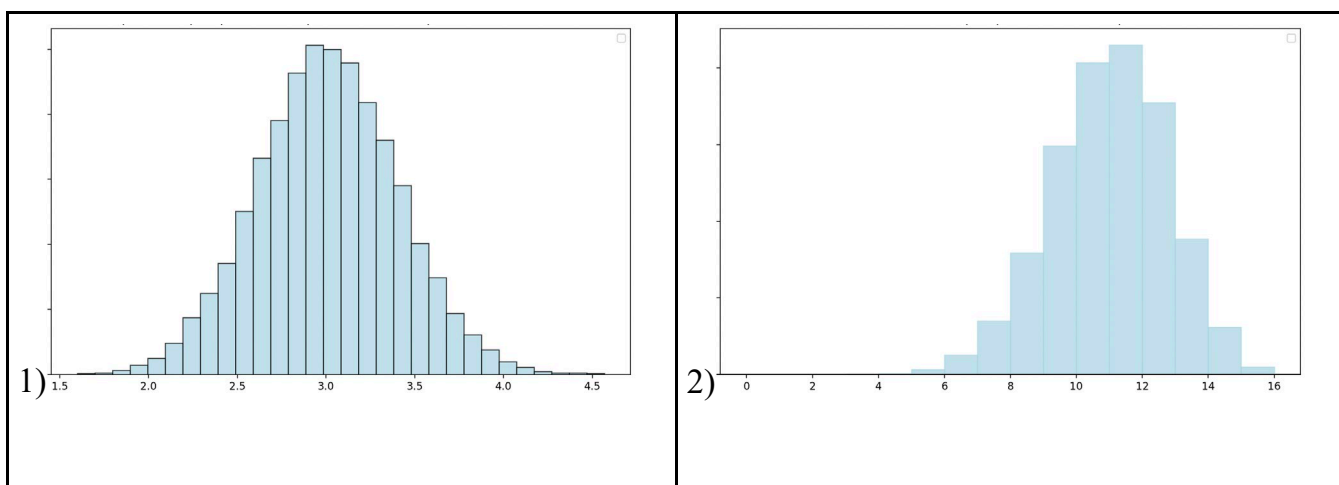
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

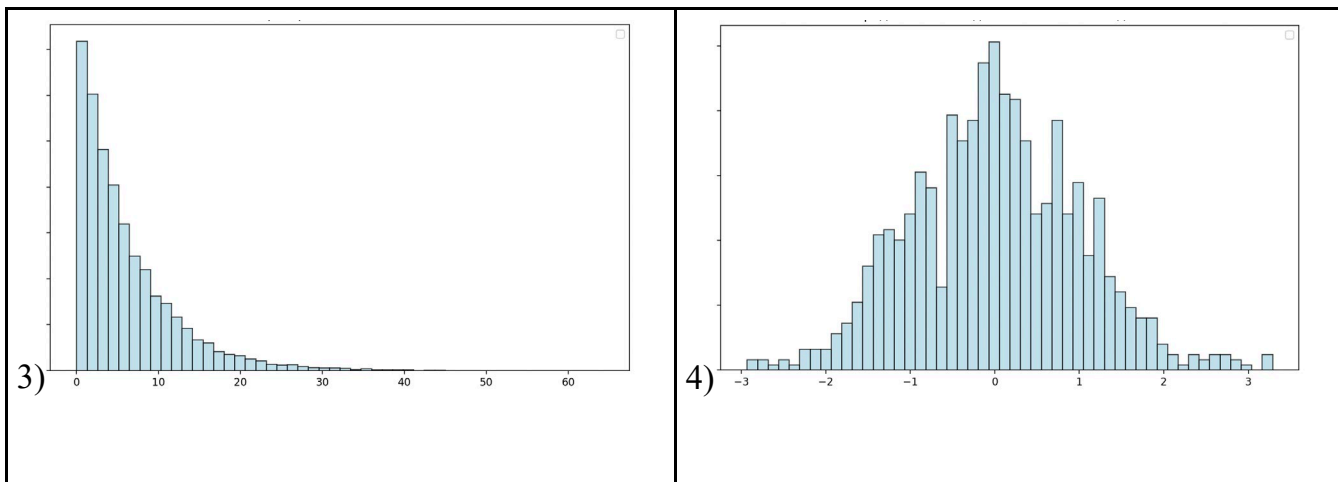
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Распределение времени до следующего сообщения в чат: среднее время, которое проходит между отправкой двух сообщений в онлайн-чате, равно 6 минут.

Б) Оценка среднего уровня удовлетворенности клиентов в небольшой компании: Применяется при оценке среднего уровня удовлетворенности клиентов, когда количество ответов ограничено.

В) Распределение уровня сахара в крови у здоровых людей: При анализе уровня сахара в крови у здоровых людей, значения варьируются, с большинством людей, имеющих средний уровень глюкозы.

Г) Распределение числа правильно решенных задач из 15 попыток, при вероятности успеха 0,7

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая проверяет, является ли введенная строка палиндромом.

Формат вывода: Строка <введенная строка> является палиндромом: <Да/Нет>.

Пример:

Ввод:

radar

Вывод:

Строка radar является палиндромом: Да.

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица products (id, name, price, category, stock_quantity, created_at), где name -наименование продукта, price- цена, category - категория продукта, stock_quantity - количество товара на складе, created_at - дата добавления продукта в систему (ГГГГ-ММ-ДД).

Задача: Найдите среднюю стоимость товаров в каждой категории.

Формат вывода: Категория, средняя цена

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (library_books.csv), содержащий информацию о книгах в библиотеке. Этот набор данных включает следующие столбцы:

- * Название книги
- * Автор
- * Год издания
- * Жанр
- * Количество выданных за последний год
- * Общее количество экземпляров

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Определите автора, чьи книги не были выданы ни разу за последний год.
2. Постройте круговую диаграмму, показывающую распределение количества экземпляров книг для пяти самых **не** популярных авторов. Убедитесь, что график имеет название и подписанные оси.

Формат вывода:

Автор, чьи книги были выданы ни разу за последний год: «Имя автора».

Решение задачи и полученные файлы сохраните в папке **задача_5**.

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Код пытается обучить логистическую регрессию, но выдает ошибку при вычислении метрик. Исправьте ошибку.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
from sklearn.linear_model import LogisticRegression

X = [[1, 2], [3, 4], [5, 6]]
y = [0, 1, 0]

model = LogisticRegression()
model.fit(X, y)
print("Accuracy:", model.score(X.reshape(-1, 1), y))
```

Код (решение) запишете в файл `task_6.py` или `задача_6.txt`

Задача 7.

Вы проектируете базу данных для управления бронированием билетов в авиакомпании.

Условия задачи:

- **Рейсы:** Каждый рейс имеет уникальный идентификатор, маршрут, дату и время вылета, самолет.
- **Самолеты:** Каждый самолет имеет уникальный идентификатор, модель и количество мест.
- **Билеты:** Каждый билет имеет уникальный идентификатор, дату бронирования и ссылку на рейс и пассажира.
- **Пассажир:** Каждый пассажир имеет уникальный идентификатор, имя, фамилия, контактные данные.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о рейсах, самолетах и билетах (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка всех пассажиров, которые летали рейсом 'RC789'.
- Поля для вывода: Имя, Фамилия

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt

Вариант 16

Ответы к задачам, записанные в файлы, необходимо сохранять в папку

Потенциал_ИТ_ФамилияИО_Вариант_16

Задача 1.

В группе из 17 человек нужно выбрать 4 для участия в соревнованиях. Сколько различных способов существует для выбора команды?

В ответ запишите развернутое решение. Для обозначения сочетаний воспользуйтесь следующей записью - $C(n, k)$.

Запишите свое решение в файл: задача_1.txt.

Задача 2.

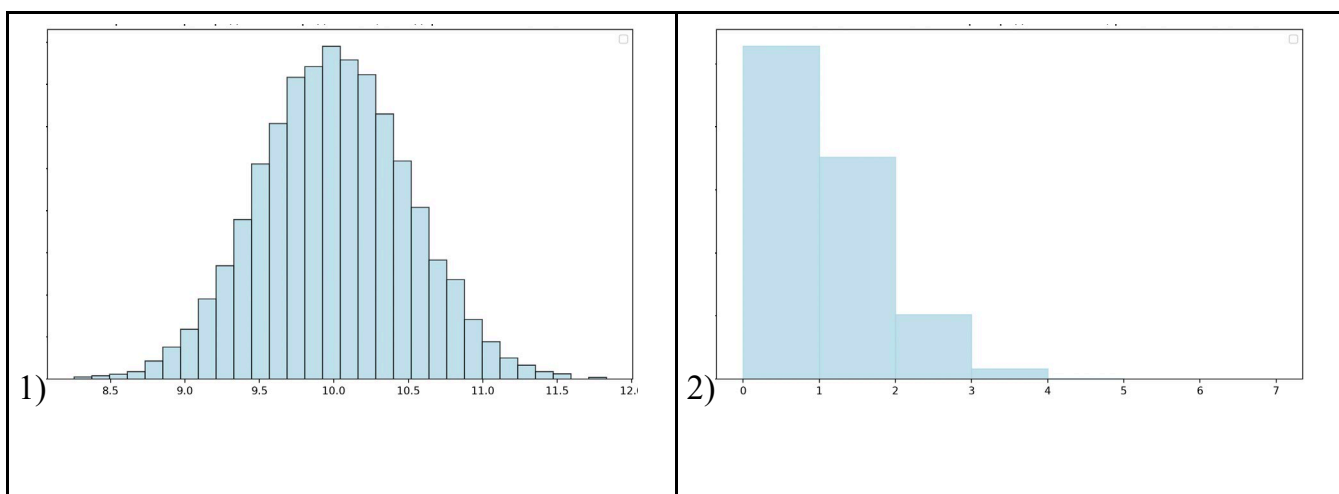
Перед вами четыре гистограммы, каждая из которых отображает распределение данных из одной из четырех выборок, описанных ниже. На гистограммах:

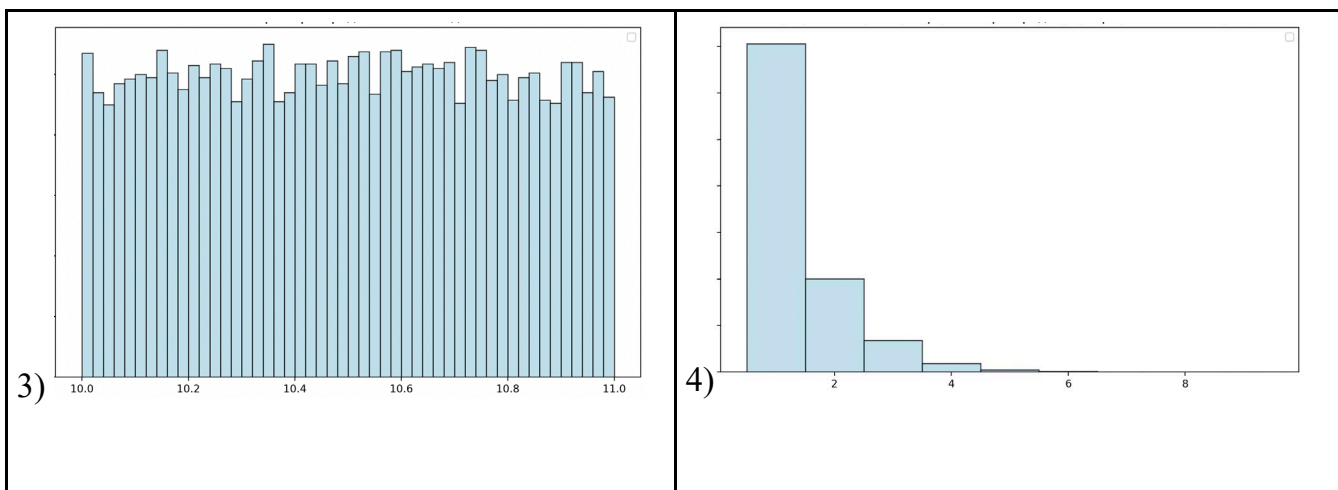
- **Ось абсцисс** показывает интервалы значений.
- **Ось ординат** отражает частоту.

Ваша задача — сравнить эти гистограммы с описаниями выборок и определить, какая из них соответствует каждому графику.

Свой ответ обоснуйте.

Запишите свое решение в файл : задача_2.txt.





А) Время прибытия поезда на станцию: Время прибытия поезда варьируется равномерно в интервале от 10:00 до 11:00 утра.

Б) Распределение числа совпадений при угадывании 6 чисел в лотерее, где вероятность совпадения одного числа составляет 0,1.

В) Распределение времени пробега 100 метров среди спортсменов: В большой выборке спортсменов время пробега 100 метров может различаться, с большинством показывающих средние результаты, и немногими — быстрее или медленнее среднего.

Г) Распределение описывает количество попыток при игре в видеоигру до первого достижения уровня: Число попыток, необходимых для достижения определенного уровня в игре.

Задача 3.

Ввод: стандартный ввод

Вывод: стандартный вывод

Напишите программу, которая находит количество уникальных букв (числа не считаются) в строке, введенной пользователем.

Формат вывода: `f"Количество уникальных символов: {count_unique}"`

Запишите свое решение в файл: task_3.py

Задача 4.

Вам дана таблица products (id, name, price, category, stock_quantity, created_at), где name -наименование продукта, price- цена, category - категория продукта, stock_quantity - количество товара на складе, created_at - дата добавления продукта в систему (ГГГГ-ММ-ДД).

Задача: Найти среднюю стоимость товаров из категории 'крупы'

Формат вывода: Средняя цена

Запишите свое решение в файл: задача_4.txt.

Задача 5.

Вам предоставлен набор данных (library_books.csv), содержащий информацию о книгах в библиотеке. Этот набор данных включает следующие столбцы:

- * Название книги
- * Автор
- * Год издания
- * Жанр
- * Количество выданных за последний год
- * Общее количество экземпляров

Используя библиотеку pandas, проведите анализ данных и предоставьте отчет по следующим критериям:

1. Найдите жанр книги, который имеет наибольшее количество экземпляров.
2. Постройте круговую диаграмму, отображающую соотношение количества экземпляров в разных жанрах.

Формат вывода:

Жанр, который имеет наибольшее количество экземпляров: X

Решение задачи и полученные файлы сохраните в папке задача_5 .

Содержимое папки оформить следующим образом:

Пункт задачи	Описание файла	Кодировка файла	Название файла
--------------	----------------	-----------------	----------------

1	В файл необходимо поместить вывод программы в указанном формате	UTF-8	задача_5_1.txt
2	График, полученный в ходе решения задачи		задача_5_2.png (.jpg)
Общий файл с программой	Полный код программы с комментариями	UTF-8	task_5.py

Задача 6.

Программа использует Pandas для вычисления среднего значения по столбцу, но выдает ошибку. Найдите и исправьте ошибку.

В качестве ответа запишите исправленный код. Аргументируйте свой ответ.

```
import pandas as pd

data = {'A': [1, 2, 3, 4]}
df = pd.DataFrame(data)

mean_value = df[1].mean()
```

Код (решение) запишете в файл task_6.py или задача_6.txt

Задача 7.

Вы проектируете базу данных для платформы онлайн-обучения.

Условия задачи:

- **Курсы:** Каждый курс имеет уникальный идентификатор, название, дату начала и стоимость. Каждый курс может вести один или несколько преподавателей.

- **Преподаватели:** Каждый преподаватель имеет уникальный идентификатор, имя, фамилию и квалификацию.

• **Студенты:** Каждый студент имеет уникальный идентификатор, имя, фамилию, и дату регистрации. Один студент может записаться на несколько курсов.

Задания:

1. Нормализация данных:

- Определите таблицы, которые необходимо создать для хранения информации о курсах, преподавателях и студентах. (опишите текстом структуру БД).
- Опишите, как вы будете устранять избыточность данных.

2. Создание структуры базы данных:

- Напишите SQL-код для создания таблиц в PostgreSQL с использованием подходящих типов данных, первичных и внешних ключей.
- Обеспечьте поддержку целостности данных с помощью ограничений.

3. Запросы:

- Напишите SQL-запрос для получения списка всех студентов с указанием количества их курсов.
- Поля для вывода: Имя студента, Фамилия студента, Количество предметов.

Решение задачи и полученные файлы сохраните в папке задача_7

Файлы в папке оформите следующим образом:

Пункт задачи	Что должно быть в файле	Кодировка файла	Название файла
Нормализация данных	В файле опишите структуру БД	UTF-8	задача_7_1.txt
Создание структуры БД	В файл добавьте SQL запрос для создания таблиц	UTF-8	задача_7_2.txt
Запросы	Добавьте SQL-запрос	UTF-8	задача_7_3.txt