**5 Лабораторная работа. Работа с файлами**

**Открытие файла**

Открытие файла осуществляется функцией open, которой нужно передать два параметра. Первый параметр — строка, задающая имя открываемого файла. Второй параметр — строка, указывающая режим открытия файла.

Существует три режима открытия файлов:

| **Режим** | **Описание** |
| --- | --- |
| "r" (read) | Файл открывается для чтения данных. |
| "w" (write) | Файл открываетсяна запись, при этом содержимое файла очищается. |
| "a" (append) | Файл открывается для добавления данных в конец файла. |

Если второй параметр не задан, то считается, что файл открывается в режиме чтения.

Функция open возвращает ссылку на **файловый объект**, которую нужно записать в переменную, чтобы потом через данный объект работать с этим файлом. Например:

input = open('input.txt', 'r')

output = open('output.txt', 'w')

Здесь открыто два файла (один на чтение, другой на запись) и создано два связанных с ними объекта.

[**Чтение данных из файла**](http://cs.mipt.ru/python/lessons/lab3.html#toc-entry-11)

Для файла, открытого на чтение данных, можно использовать несколько методов для считывания данных. Например, readline(), readlines(), read().

Метод readline() считывает одну строку из файла (до символа конца строки 'n', возвращается считанная строка вместе с символом 'n'). Если считывание не было успешно (достигнут конец файла), то возвращается пустая строка. Для удаления символа 'n' из конца файла удобно использовать метод строки rstrip(). Например, s = s.rstrip().

Метод readlines() считывает все строки из файла и возвращает список из всех считанных строк (одна строка — один элемент списка). При этом символы 'n' остаются в концах строк.

Метод read() считывает все содержимое из файла и возвращает строку, которая может содержать символы 'n'. Если методу read передать целочисленный параметр, то будет считано не более заданного количества байт. Например, считывать файл побайтово можно при помощи метода read(1).

[**Запись данных в файл**](http://cs.mipt.ru/python/lessons/lab3.html#toc-entry-12)

Данные выводятся в файл при помощи метода write(), которому в качестве параметра передается одна строка. Этот метод не выводит символ конца строки 'n' (как это делает функция print() при стандартном выводе), поэтому для перехода на новую строку в файле необходимо явно вывести символ 'n'.

Выводить данные в файл можно и при помощи print(), если передать функции еще один именованный параметр file. Например:

output = open('output.txt', 'w')

print(a, b, c, file=output)

[**Закрытие файла**](http://cs.mipt.ru/python/lessons/lab3.html#toc-entry-13)

После окончания работы с файлом необходимо закрыть его при помощи метода close().

Чтобы не забыть это сделать можно воспользоваться *менеджером контекста* with:

**with** open('input.txt') **as** file:

[**Примеры работы с файлами**](http://cs.mipt.ru/python/lessons/lab3.html#toc-entry-14)

Следующая программа считывает все содержимое файла input.txt, записывает его в переменную s, а затем выводит ее в файл output.txt:

inp = open('input.txt', 'r')

out = open('output.txt', 'w')

s = inp.read()

out.write(s)

inp.close()

out.close()

Для простого считывания содержимого файла можно использовать то, что сам файл является итерируемым по строкам объектом:

**with** open('input.txt') **as** file:

**for** line **in** file:

print('line: "', line, '"')

**Задание на лабораторную работу:**

1. В файле input.txt записаны числа через пробел. Просуммировать первые пять чисел и ответ записать в файл output.txt
2. Из всевозможных комбинаций двух чисел, записанных в файле input.txt, вычислить произведения.
3. Посмотрите на шрифт для написания почтового индекса на конвертах:



Воспользуйтесь списками кортежей, чтобы задать рисование Черепашкой таких цифр. Нарисуйте на экране индекс (индивидуальное значение возьмите у преподавателя).

Перенесите описание способа рисования почтовых цифр (списки движений) в файл. Пусть черепаха считывает "шрифт" из файла.

1. Для выполнения данного упражнения вам понадобится файл titanic.csv

Из входных данных необходимо выделить все записи, где пассажиром является незамужняя женщина. Все полученные данные необходимо сохранить в новый файл. Алгоритм должен быть реализован в виде функции с именем `prepare\_data`, которая на вход принимает имена входного и выходного файлов."

1. Входные данные лежат в файле с названием titanic.csv.

Посчитайте средний возраст выживших незамужних женщин и распечатайте его. Алгоритм должен быть реализован в виде функции с именем `get\_mean\_age`, которая на вход принимает имя входного файла и возвращает средний возраст."

1. Посчитайте долю пропущенных ячеек по каждому столбцу в файле с названием titanic.csv. Алгоритм должен быть реализован в виде функции, которая на вход принимает имя входного файла и возвращает словарь, в котором ключи – это названия столбцов, а значения доля пропущенных ячеек в каждом столбце.
2. Из входных данных выделите только тех пассажиров, по которым есть полная информация (не пропущена ни одна ячейка) и запишите их в отдельный файл. Алгоритм должен быть реализован в виде функции, которая на вход принимает имя входного файла и возвращает количество пассажиров со всей информацией.