Министерство науки и высшего образования Российской Федерации Федеральное государственное автономное образовательное учреждение высшего образования «СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт цифрового развития Кафедра инфокоммуникаций

ОТЧЕТ

По лабораторной работе №2 Дисциплины «Анализ данных»

Выполнил:

Пустяков Андрей Сергеевич

2 курс, группа ИВТ-б-о-22-1,

09.03.01 «Информатика и вычислительная техника (профиль) «Программное обеспечение средств вычислительной техники и автоматизированных систем», очная форма обучения

(подпись)

Руководитель практики:

Воронкин Р. А. кандидат технических наук, доцент, доцент кафедры инфокоммуникаций

(подпись)

Тема: Работа с данными формата JSON в языке Python.

Цель: приобрести навыки по работе с данными формата JSON с помощью языка программирования Python версии 3.х.

Ход работы:

Создание общедоступного репозитория на «GitHub», клонирование репозитория, редактирование файла «.gitignore», организация репозитория согласно модели ветвления «git-flow» (рис. 1).

```
C:\Users\Andrey\Desktop\Анализ данных\2 лаба\Analysis_data_laboratory_work_2>git flow init

Which branch should be used for bringing forth production releases?
- main
Branch name for production releases: [main]
Branch name for "next release" development: [develop]

How to name your supporting branch prefixes?
Feature branches? [feature/]
Bugfix branches? [bugfix/]
Release branches? [release/]
Hotfix branches? [nelease/]
Hotfix branches? [support/]
Version tag prefix? []
Hooks and filters directory? [C:/Users/Andrey/Desktop/Анализ данных/2 лаба/Analysis_data_laboratory_work_2/.git/hooks]
```

Рисунок 1 – Организация модели ветвления «git-flow».

Проработка примеров лабораторной работы:

Пример 1.

Необходимо создать программу, в которой использовать словарь, содержащий информацию о фамилии и инициалах работника, название занимаемой должности, год поступления на работу. В программе должен быть организован ввод данных в список, состоящий из словарей заданной структуры, записи должны быть упорядочены по алфавиту. Должен быть организован вывод на дисплей фамилии работников, чей стаж работы в организации превышает значение, введенное с клавиатуры, причем, если таких работников нет, то должно быть выведено соответствующее сообщение. Необходимо добавить возможность сохранения списка в файл формата JSON и чтения данных из JSON.

Код программы решения данной задачи и результаты работы программы с различными исходными данными (рис. 2, 3, 4).

```
year = int(input("Год поступления? "))
   if staff:
                    idx, worker.get('name', ''),
def select workers(staff, period):
```

```
today = date.today()
for employee in staff:
    if today.year - employee.get('year', today.year) >= period:
        result.append(employee)
   json.dump(staff, fout, ensure ascii=False, indent=4)
workers = []
while True:
        workers.append(worker)
        display workers(workers)
        parts = command.split(maxsplit=1)
```

```
selected = select_workers(workers, period)
parts = command.split(maxsplit=1)
file_name = parts[1]
parts = command.split(maxsplit=1)
file name = parts[1]
workers = load workers(file name)
```

Рисунок 2 – Код программы примера 1

Рисунок 3 – Результаты работы программы без подгруздки данных

```
      C:\Users\Andrey\AppData\Local\Programs\Python\Python39\python.exe
      "C:\Users\Andrey\Decorporal C:\Users\Andrey\Decorporal C:\Users\Andrey
```

Рисунок 4 — Результаты работы программы с подгрузкой данных из файла Выполнение индивидуальных заданий:

Задание 1.

Необходимо для своего варианта лабораторной работы 2.8 дополнительно реализовать сохранение и чтение данных из файлов формата JSON (генерируемые файлы не должны находиться в репозитории лабораторной работы).

Необходимо использовать словарь, содержащий следующие ключи: название пункта назначения; номер поезда; время отправления. Написать программу, выполняющую следующие действия: ввод с клавиатуры данных в список, состоящий из словарей заданной структуры; записи должны быть упорядочены по времени отправления поезда; вывод на экран информации о поездах, направляющихся в пункт, название которого введено с клавиатуры; если таких поездов нет, выдать на дисплей соответствующее сообщение (Вариант 26 (7), работа 2.8).

Код программы для решения данной задачи, для избежания сохранения файлов в папке репозитория была осуществлена смена рабочей директории в программе (рис. 5).

```
if staff:
        line = '+-{}-+-{}-+-{}-+-{}-+--{}-+'.format(
        for idx, points in enumerate(staff, 1):
                    idx, points.get('departure point', ''),
                    points.get('number train', ''),
                    points.get('destination', ''))
def select trains(staff, point user):
            result.append(train)
os.chdir("C:\\Users\\Andrey\\Desktop\\Анализ данных\\2 лаба\\Kaliningrad")
        json.dump(staff, fout, ensure ascii=False, indent=4)
```

```
def load trains(file name):
os.chdir("C:\\Users\\Andrey\\Desktop\\Анализ данных\\2 лаба\\Kaliningrad")
            trains.append(train)
            display trains(trains)
            point comand = command.split(maxsplit=1)
            parts = command.split(maxsplit=1)
            file_name = parts[1]
            save trains(file name, trains)
            parts = command.split(maxsplit=1)
            file name = parts[1]
            trains = load trains(file name)
```

```
elif command == 'help':
    # Вывести справку о работе с программой.
    print("Список команд:\n")
    print("add - добавить поезд;")
    print("list - вывести список поездов;")
    print("select <cтанция> - запросить поезда направляющиеся в

пункт;")
    print("help - отобразить справку;")
    print("load - загрузить данные из файла;")
    print("save - сохранить данные в файл;")
    print("exit - завершить работу с программой.")

else:
    print(f"Hеизвестная команда {command}", file=sys.stderr)

if __name__ == '__main__':
    main()
```

Рисунок 5 – Код программы решения данной задачи

Результаты работы программы с загрузкой списка из файла формата JSON (рис. 6).

```
| 1 | Айвазовская | 001 | 6:00 | Калининград Ю. |
| 2 | Держинская | 002 | 6:10 | Калининград Ю. |
| 3 | Киевская | 003 | 6:20 | Калининград Ю. |
| 4 | Западная | 004 | 6:30 | Калининград Ю. |
| 5 | Калининград С. | 005 | 6:40 | Калининград Ю. |
| ***>>> select Калиниерад С.
Список станций пуст.
>>> select Калининград С.

| № | Пункт отправления | Номер поезда | Время отправления | Пункт назначения |
| 1 | Сельма | 006 | 6:50 | Калининград С. |
| 2 | Кутузово | 007 | 7:00 | Калининград С. |
| 3 | Калининград Ю. | 008 | 7:10 | Калининград С. |
| ***>>> exit
```

Рисунок 6 – Результаты работы программы

Содержимое файла «Kaliningrad.json» (рис. 7).

Рисунок 7 – Файл «Kaliningrad.json»

Задание повышенной сложности:

Необходимо организовать валидацию загружаемых данных из файла формата JSON с помощью спецификации JSON Schema.

Для того, чтобы производить валидацию файлов необходимо установить виртуальное окружение «jsonschema» (рис. 8).

```
(base) C:\Users\Andrey>conda create --name lab_2
Retrieving notices: ...working... done
Collecting package metadata (current_repodata.json): done
Solving environment: done
==> WARNING: A newer version of conda exists. <==
 current version: 23.7.4
  latest version: 24.3.0
Please update conda by running
    $ conda update -n base -c defaults conda
Or to minimize the number of packages updated during conda update use
     conda install conda=24.3.0
## Package Plan ##
  environment location: C:\Users\Andrey\anaconda3\envs\lab_2
Proceed ([y]/n)? y
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
 To activate this environment, use
      $ conda activate lab_2
  To deactivate an active environment, use
      $ conda deactivate
```

Рисунок 8 – Создание нового виртуального окружения

Установка необходимых пакетов для виртуального окружения (рис. 9).

```
(base) C:\Users\Andrey>conda activate lab_2

_(lab_2) C:\Users\Andrey>conda install jsonschema
orCollecting package metadata (current_repodata.json): done
Solving environment: done
```

Рисунок 9 – Установка «jsonschema»

Добавление валидации подгружаемых файлов в функцию подгрузки файлов (рис. 10).

```
print(f"Произошла ошибка валидации! {e.message}")
```

Рисунок 10 – Функция загрузки JSON файлов

Результаты работы данной функции для подходящего и неподходящего файлов (рис. 11).

```
C:\Users\Andrey\AppData\Local\Programs\Python\Python39\python.exe C:\Users\Andrey\
>>> load kaliningrad_2
Произошла ошибка валидации! 'departure_point' is a required property
>>> load kaliningrad
JSON-файл прошел валидацию по заданной схеме.
>>>
```

Рисунок 11 – Валидация загружаемых файлов

Ответы на контрольные вопросы:

1. Для чего используется JSON?

JSON — это стандарт обмена данными. Он позволяет легко сериализовать и десериализовать объекты. Стандарт часто применяют, когда разрабатывают API и веб-приложения.

2. Какие типы значений используются в JSON?

В качестве значений в JSON могут быть использованы: запись — это неупорядоченное множество пар «ключ: значение», заключённое в фигурные скобки «{ }». Ключ описывается строкой, между ним и значением стоит символ «:». Пары ключ-значение отделяются друг от друга запятыми. Список (одномерный) — это упорядоченное множество значений. Список заключается в квадратные скобки «[]». Значения разделяются запятыми. Число (целое или вещественное). Литералы true (логическое значение «истина»), false (логическое значение «ложь») и null. Строка — это упорядоченное множество из нуля или более символов юникода, заключённое в двойные кавычки.

3. Как организована работа со сложными данными в JSON?

JSON позволяет организовать сложные структуры данных, такие как списки и вложенные словари (объекты). В JSON можно хранить разные типы данных, включая числа, строки, логические значения, массивы и объекты. Для организации сложных данных в JSON используются вложенные объекты и списки, позволяя создавать структуры данных любой сложности.

4. Самостоятельно ознакомьтесь с форматом данных JSON5? В чем отличие этого формата от формата данных JSON?

JSON5 — это расширение формата данных JSON, разработанное для улучшения читаемости и удобства записи JSON-данных. Отличие JSON5 от обычного JSON включает в себя дополнительные возможности, такие как использование комментариев, разделителей ключей и значений, а также возможность использования одиночных кавычек вместо двойных. JSON5

является более гибким и читаемым форматом для записи данных, но не является стандартом и не поддерживается всеми JSON-парсерами.

5. Какие средства языка программирования Python могут быть использованы для работы с данными в формате JSON5?

Для работы с данными в формате JSON5 на Python, можно использовать парсеры, поддерживающие JSON5, такие как demjson. Однако, JSON5 не является стандартом, поэтому поддержка может быть ограничена.

6. Какие средства предоставляет язык Python для сериализации данных в формате JSON?

Для сериализации данных в формат JSON в Python можно использовать модуль json. Он предоставляет функции json.dump() и json.dumps(), а также класс json.JSONEncoder, который может быть настроен для сериализации данных в формат JSON.

7. В чем отличие функций json.dump() и json.dumps()?

«json.dump()» записывает данные в файл. Вы используете его, когда хотите сохранить данные в файле. json.dumps() превращает данные в строку. Вы используете его, когда хотите получить данные в виде строки для дальнейшей обработки, но не сохранять их в файле.

8. Какие средства предоставляет язык Python для десериализации данных из формата JSON?

Для десериализации данных из формата JSON в Python используется модуль json, предоставляющий функции json.load() и json.loads().

9. Какие средства необходимо использовать для работы с данными формата JSON, содержащими кирилицу?

Для работы с данными JSON, содержащими кириллицу, важно убедиться, что данные правильно кодируются и декодируются. Обычно это не вызывает проблем, поскольку JSON поддерживает Unicode, включая кириллические символы. Однако, при чтении и записи JSON-файлов, убедитесь, что правильно установлена кодировка (например, utf-8) для текстовых данных.

10. Самостоятельно ознакомьтесь со спецификацией JSON Schema? Что такое схема данных?

JSON Schema - это спецификация, которая описывает формат данных JSON и правила их валидации. С помощью JSON Schema можно определить структуру, типы данных и ограничения для JSON-данных. JSON Schema используется для проверки соответствия данных определенным правилам. Это полезно, например, при валидации данных, получаемых из внешних источников. JSON Schema не является частью стандартной библиотеки Руthon, но существуют библиотеки и инструменты, поддерживающие JSON Schema, которые могут использоваться в Руthon.

Вывод: в ходе выполнения лабораторной работы были приобретены навыки по работе с JSON-файлами, а также была изучена спецификация «JSON Schema», позволяющая проводить валидацию подгружаемых данных.