

## БЛОК I. РАЗМИНКА

**I.1** Файл «**ozone.csv**» предоставляет таблицы с данными по общим объемам озона в октябре в единицах Добсона<sup>1</sup> и по двум компонентам хлорфторуглерода F11 и F12 в числе частей на триллион по объему (pptv) за период 1957–1984 гг.<sup>2</sup>.

Прочитать и выполнить парсинг этих данных и изобразить их на наиболее подходящей диаграмме.

**I.2 На странице**<sup>3</sup> Википедия предоставляет список значений распространенности элементов на Солнце и в Солнечной системе в виде таблицы HTML (наряду с другими подобными данными). Использовать метод *read\_html* из библиотеки *pandas* для чтения и парсинга данных Кайе и Лаби (Kaye and Laby) (столбец с заголовком «Y1») и создать столбиковую диаграмму (гистограмму), демонстрирующую правило Оддо-Гаркинса: элементы с четными атомными номерами более распространены, чем соседние с ними элементы с нечетными атомными номерами.

**I.3** Диаграмма Герцшпрунга–Рассела (или Рассела) классифицирует звезды на точечной диаграмме: каждая звезда представлена точкой, координатой *x* которой является активная температура, а координатой *y* – светимость (звездная величина), мера электромагнитного излучения звезды. Файл «**hygdata\_v3-abridged.csv**» предоставляет данные базы данных HYG-database<sup>4</sup>, в которой представлены данные о 119 614 звездах. Прочитать эти данные с помощью методов библиотеки *pandas* и создать диаграмму Герцшпрунга–Рассела. Столбец светимости идентифицируется как '**lum**' в заголовке, а температуру звезды можно вычислить по ее показателю цвета (также обозначаемому как (B-V) и определяемому по столбцу с меткой '**ci**'), воспользовавшись формулой Баллестероса:

$$\frac{T}{K} = 4600 \left( \frac{1}{0.92(B - V) + 1.7} + \frac{1}{0.92(B - V) + 0.62} \right)$$

Следует отметить, что светимость лучше всего отображается на логарифмической шкале, а ось температуры обычно изображается в обратном направлении (с уменьшением температуры слева направо на диаграмме).

**I.4** Transport for London (TfL) – это публично-правовая корпорация, явля-

<sup>1</sup>Единица Добсона (Dobson unit) равна слою чистого озона 10 мкм (0.01 мм) при стандартных давлении и температуре, образующегося из общего объема озона в атмосфере над какой-либо областью поверхности Земли.

<sup>2</sup>См. Farman et al., Nature 315, 207 (1985)

<sup>3</sup>[https://en.wikipedia.org/wiki/Abundances\\_of\\_the\\_elements\\_\(data\\_page\)](https://en.wikipedia.org/wiki/Abundances_of_the_elements_(data_page))

<sup>4</sup>База данных <https://github.com/astronexus/HYG-Database>, опубликованная под защитой лицензии Creative Commons Attribution-ShareAlike.

яся функциональным подразделением администрации Лондона и отвечающая за управление транспортной системой Лондона, включая координацию работы общественного транспорта, содержание главных дорог и светофоров. В файле **«multi-year-station-entry-and-exit-figures.xls»** содержатся статистические данные об использовании сети подземного транспорта (The Tube – лондонское метро) в форме количества входящих и выходящих пассажиров за «обычный» день на каждой станции за период 2007–2017 гг.

Прочитать этот документ средствами библиотеки pandas и проанализировать его, чтобы определить:

- самую загруженную станцию в обычный день в 2017 г.;
- станцию с наибольшим приростом пассажиропотока за период 2007–2017 гг.;
- станцию с наибольшей относительной разностью между количеством пассажиров в рабочие дни и в обычное воскресенье в 2017 г.

**I.5** В файле **«CO2-transitions.par»** содержится база данных HITRAN <sup>5</sup>, которая предоставляет список значений интенсивности молекулярных линий для моделирования лучистого теплообмена в атмосферах планет (астрофизика молекулярных линий). Собственный формат этой базы данных представлен в виде записей длиной 160 символов с фиксированной шириной полей.

Использовать pandas для чтения файла **«CO2-transitions.par»**. Данные представлены в виде полей фиксированной длины. Нужные поля содержатся в колонках 4 – 15 (волновые числа,  $\tilde{\nu}$ ,  $\text{см}^{-1}$ ) и колонки 16 – 25 (интенсивность спектральных (молекулярных) линий, S, при  $T = 296\text{K}$ ,  $\text{см}^{-1}/\text{площадь молекулы (?) см}^{-2}$ )

Построить график зависимости интенсивности молекулярных линий от длины волны при лучистом теплообмене в инфракрасной области спектра ( $\lambda =$  от 1 микрона до 20 микрон, соответствующая волновым числам  $\tilde{\nu} =$  от  $500\text{см}^{-1}$  до  $10000\text{см}^{-1}$ ), где диоксид углерода ( $\text{CO}_2$ ) отвечает в значительной степени за возникновение парникового эффекта в атмосфере Земли.

---

<sup>5</sup><https://hitran.org>