**ТЕСТОВОЕ ЗАДАНИЕ ПО АНАЛИЗУ ДАННЫХ В PYTHON**

*(на выполнение предполагается дать 3 дня)*

Тестовое задание сдаётся в формате Jupyter Notebook (ipynb) или ссылкой на код на Github. Пожалуйста, убедитесь, что Ваш код содержит все необходимые, графики, комментарии и предпосылки, которые были сделаны в ходе решения.

На <https://yadi.sk/d/MuItohAzWJ_tnA> Вы сможете скачать файлы:

1. units.csv – массив данных о торговом графике режимных генерирующих единиц, публикуемых АТС (<https://www.atsenergo.ru/nreport?rname=carana_sell_units>) за период 01/10/2018 – 01/10/2019:
   * datetime;
   * unit\_id;
   * node\_id;
   * pdg;
   * upper\_limit;
2. prices.csv – массив данных о равновесных ценах в наиболее крупных узлах расчётной модели, публикуемых АТС (<https://www.atsenergo.ru/nreport?rname=big_nodes_prices_pub>) за период 01/10/2018 – 01/10/2019:
   * datetime;
   * node\_id;
   * price.
3. indices.xls – массив данных о равновесных ценах РСВ (<http://www.atsenergo.ru/results/rsv/index>) за период 01/10/2018 – 01/10/2019:
   * datetime;
   * price\_zone;
   * volume\_bid;
   * price\_bid.
4. nodes.xls – реестр узлов расчётной модели:
   * node\_id;
   * region;
5. regions – перечень субъектов РФ и их принадлежность;
   * regions;
   * federal\_district;
   * price\_zone;
   * price\_zone\_id.

В комментариях к коду, дайте развернутые ответы на следующие вопросы:

1. Придумайте как кластеризовать генерирующие единицы (unit\_id) 1-й ценовой зоны (метод).
2. Опишите, чем кластеры отличаются друг от друга.
3. Дайте своё объяснение, что могут означать кластеры?
4. Изменяются ли характеристики поведения «кластеров» в течение рассмотренного периода?