Загрузка данных и аналитка

* Загрузите файл Schools Data.xlsx
* Проведите первоначальный разведывательный анализ – какие имеются колонки? Какие значения в интересующих вас категориальных колонках?
* Воспользуйтесь файлом GDP by region.xlsx для того, чтобы добавить в имеющийся DataFrame ещё одну колонку – ВВП данного региона.
* С помощью простой визуализации (например гистограмм - .plot.hist()) или статистических показателей (среднее, медиана, квартили) попробуйте проанализировать взаимосвязь между ВВП региона и различными показателями школ – качеством образования, материальной базой и т.д.

<https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.plot.hist.html> - документация по построению гистограмм

* Попробуйте проверить какие-то ещё простые гипотезы, не связанные с ВВП (например, что количество учителей высшей категории положительно влияет на баллы ЕНТ)

Отчёты

* Напишите программу, которая автоматически разделит ваш файл на 16 файлов по регионам
* Напишите программу, которая выведет чёрный список школ по значению колонки qedu\_passed\_ent (количество набравших проходной балл ЕНТ) – например те школы, в которых qedu\_passed\_ent меньше 0.5. Программа должна выводить отдельный чёрный список для каждого региона

2.1) Выведите список школ с проблемами с материальной обеспеченностью (например, те школы, где требуется капитальный ремонт, туалеты на улицах, нет горячего питания) – конкретные условия на ваше усмотрение

3) На основании колонок eng\_informatics\_teachers, eng\_phys\_teachers, eng\_chem\_teachers и eng\_bio\_teachers составьте два отчёта:

3.1) Топ школ по количеству учителей, владеющих английским языком

3.2) Отдельные отчёты с топом школ по каждый специальности (informatics, phys, chem, bio)

4) Выведите один файл с статистикой по каждому населённому пункты. Необходимо посчитать для каждого населённого пункта: количество школ, процент школ с доступом в интернет, наличие хотя бы одной школы без доступа к питьевой воде (да или нет).

4.1) По своему усмотрению добавьте другие статистики и агрегаты в этот отчёт по населённым пунктам. Попробуйте найти в этом отчёте какие-то инсайты, проведите анализ.