Жуковский Павел, 4 курс, 12 группа

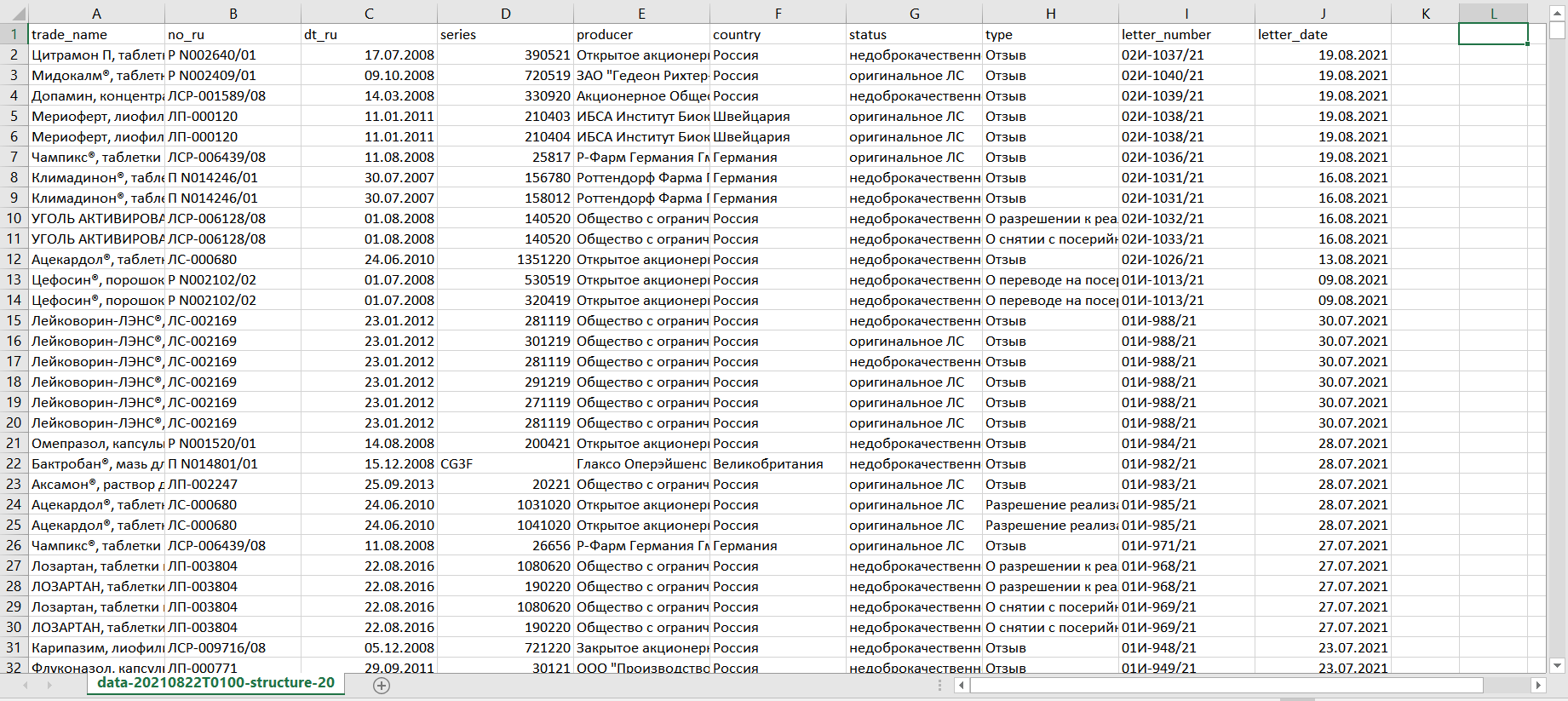
Сбор и подготовка открытых данных для анализа

Часть 1 Исследование наборов данных на портале data.gov.ru:   
**Задание**.

1 Исследование наборов данных на портале data.gov.ru:   
1.1. Выберите вариант задания (табл. 5)   
1.2. Найдите произвольный набор данных на портале data.gov.ru по тематике, указанной в выбранном варианте задания. Набор должен быть представлен в формате csv и кодировке Windows.   
1.3 Загрузите на компьютер найденный набор данных и его паспорт.   
1.4 Проведите анализ набора данных: определите количество записей и полей в наборе данных.

В качестве тематики я выбрал **Здоровье** (<http://data.gov.ru/rubriki/zdorove>)

**Копия экрана с набором данных, открытом в Microsoft Excel.**

Сведения об осуществлении государственного контроля качества лекарственных средств****

**Описание набора данных**

|  |  |
| --- | --- |
| Показатель | Значение |
| Наименование | Сведения об осуществлении государственного контроля качества лекарственных средств |
| Ссылка | <https://data.gov.ru/opendata/7710537160-drugsqualitycontrol> |
| Формат | CSV |
| Количество записей | 35708 |
| Количество полей | 10 |
| В т.ч числовых | 1 (общее поле, но в основном с числами) |
| В т.ч текстовых | 9 (включая поля с датой) |

Часть 2 Исследование наборов данных на портале data.worldbank.org:   
**Задание**.

2. Исследование наборов данных на портале data.worldbank.org:   
2.1 Выберите вариант задания (табл. 6) Найдите произвольный набор данных на портале data.worldbank.org по тематике, указанной в выбранном варианте задания.   
2.3 Загрузите на компьютер найденный набор данных в формате xls.   
2.4 На основе набора данных подготовьте выборку, содержащую значения показателя за все годы для трёх произвольно выбранных стран мира.   
2.5. На основе подготовленной выборки постройте график, иллюстрирующий изменение показателя со временем для трёх стран мира.   
2.6. Сохраните файл.

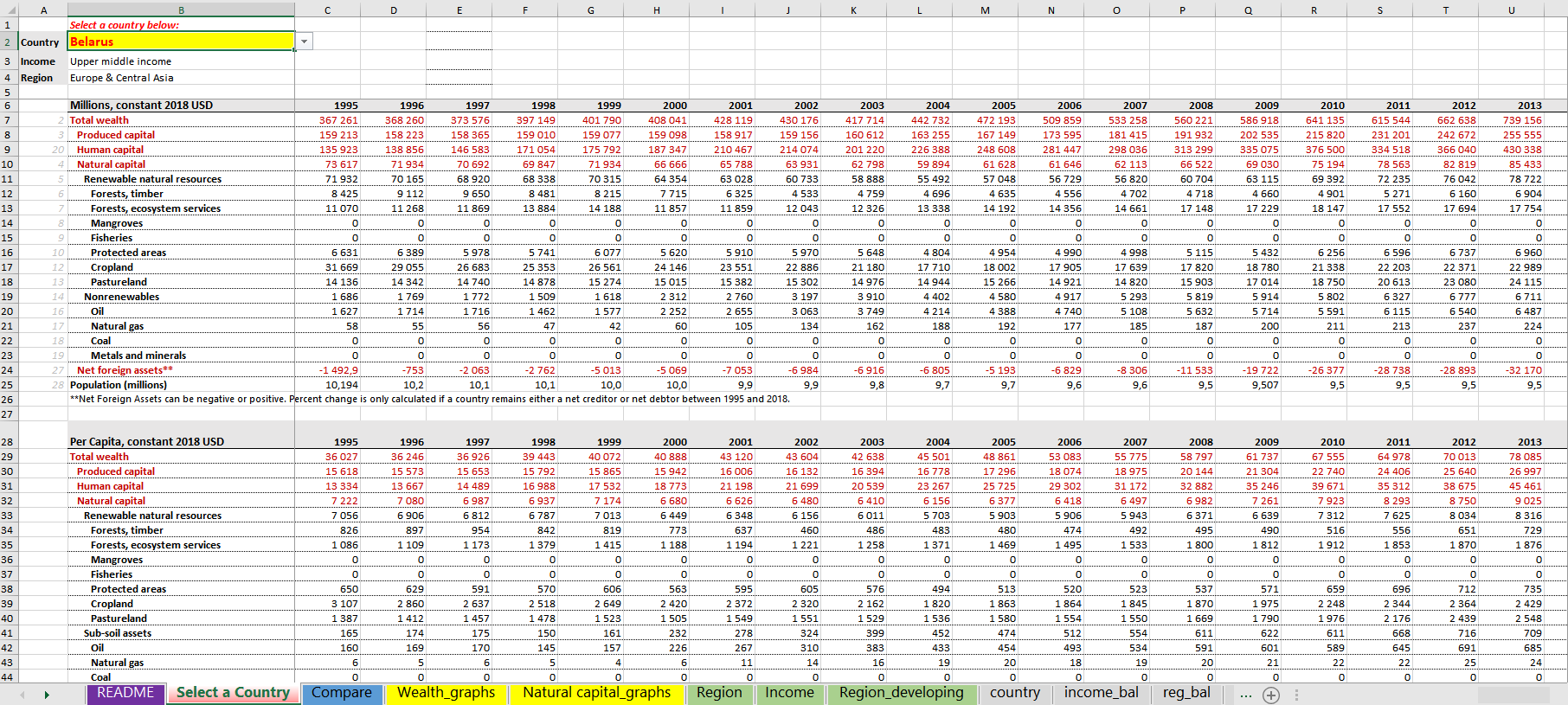
**Копия экрана с набором, содержащим данные.**

В качестве тематики я выбрал **Energy & Mining**.

В качестве набора данных я выбрал:

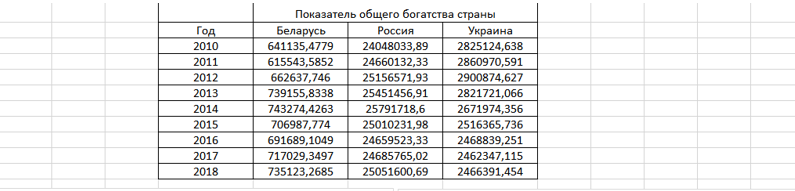
The Changing Wealth of Nations 2021 (Изменяющееся богатство наций 2021 г.)

Ссылка на ресурс: <https://datacatalog.worldbank.org/search/dataset/0042066/Wealth-Accounting>



**График изменения показателя со временем по трем произвольно выбранным странам мира**

Я взял некоторые данные из вышеуказанной таблицы и на их основе построил график зависимости показателя общего богатства от времени для стран: Беларусь, Россия и Украина. Так график выглядит не очень красиво и наглядно (из-за больших показателей России), я также построил эквивалентную диаграмму.

****

****

Часть 3 Построение автоматизированной системы сбора данных:   
**Задание.**

3. Построение системы автоматизированной системы сбора данных:   
3.1 Выберите вариант задания табл. 7.   
3.2. Найдите интернет-сайт, содержащий указанные в задании данные.   
3.3. Запустите Microsoft Excel.   
3.4. Выберите пункт «Из Веба» в меню «Данные».   
В адресной строке появившегося окна «Создание веб-запроса» наберите адрес найденной ранее веб-страницы. Выберите таблицу, содержащую искомые данные. Нажмите кнопку «Импорт».   
В появившемся окне «Импорт данных» нажмите кнопку «Свойства».

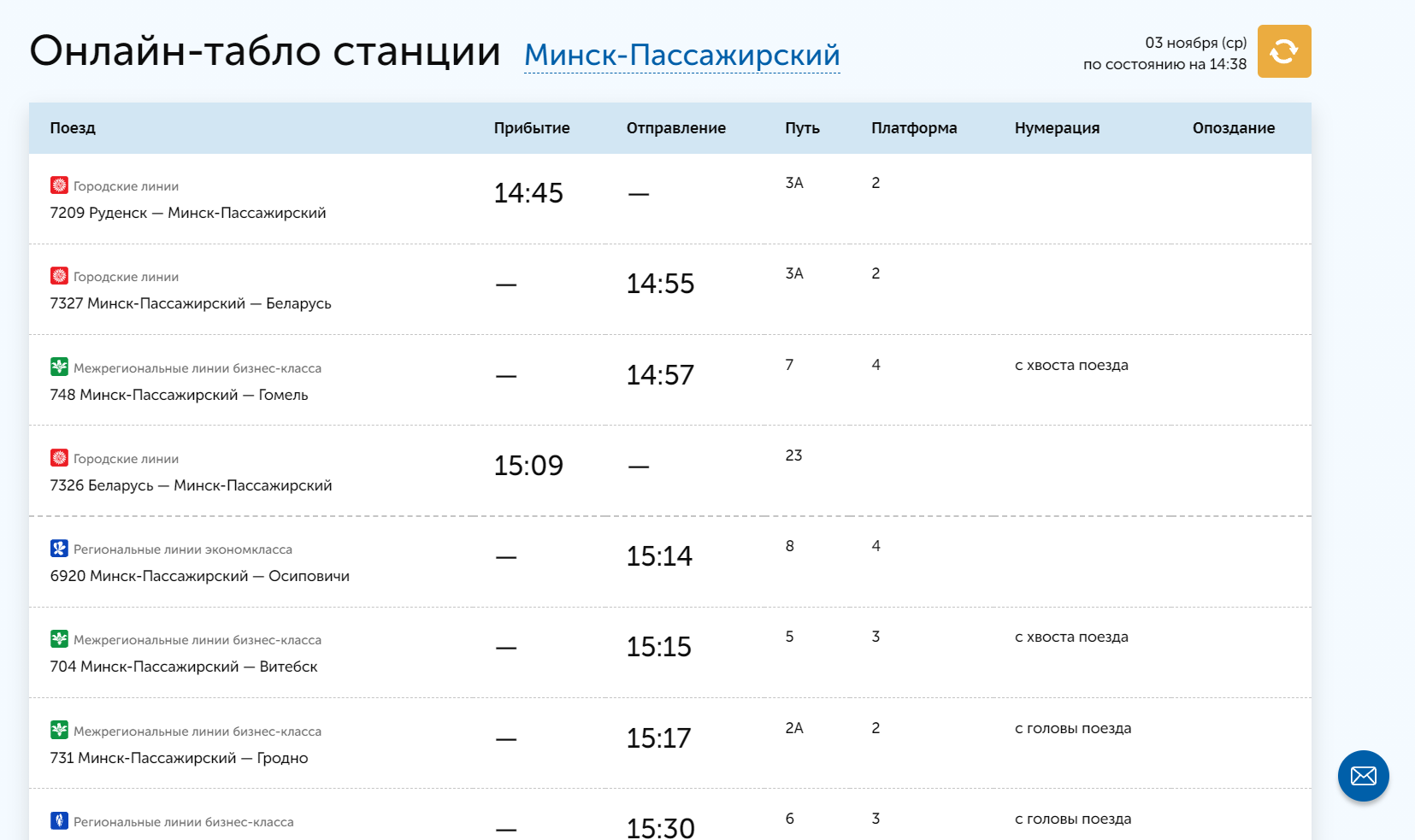
В появившемся окне «Свойства внешнего диапазона» задайте параметр «Период обновления», равный 1 минуте, параметр «Обновление при открытии файла» «Да». Нажмите кнопку «ОК».   
3.5. В окне «Импорт данных» нажмите кнопку «ОК».   
Сохраните файл

**Копия экрана с интернет-сайтом, содержащим данные.**

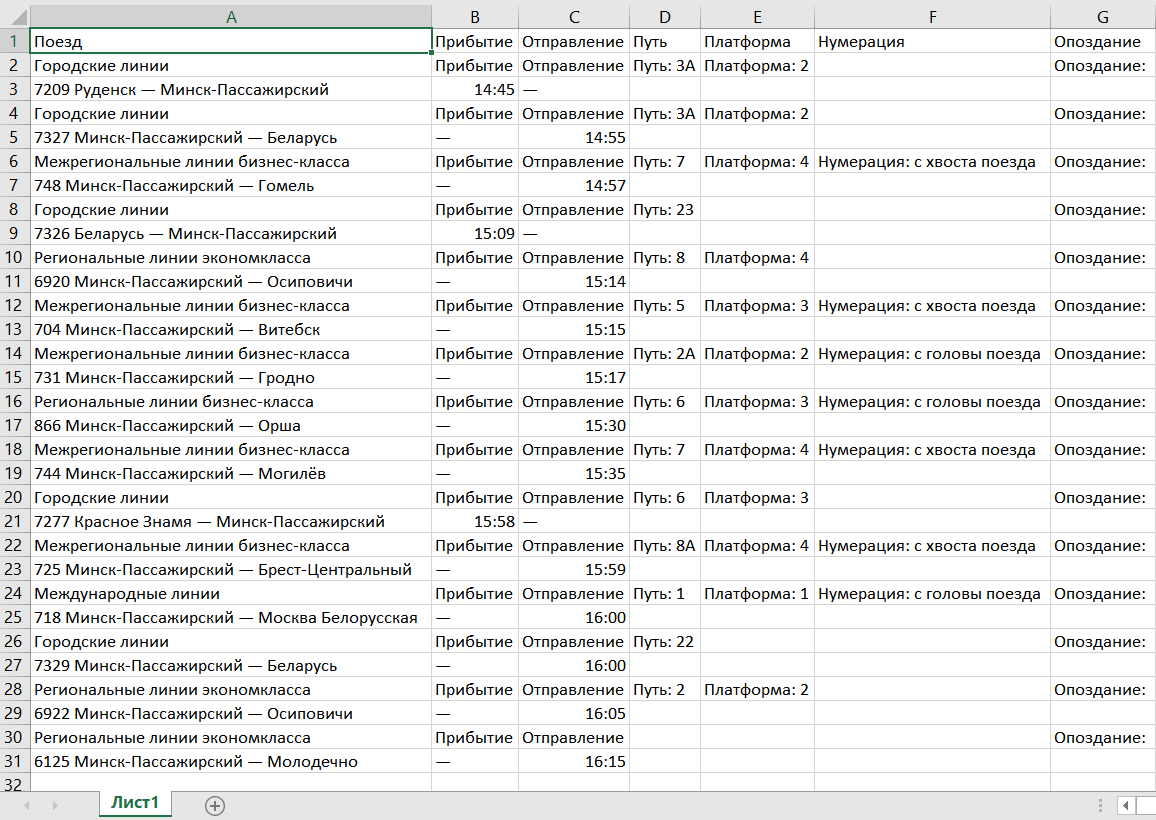
В качестве тематики я выбрал **Онлайн-табло аэропорта**

Я пробовал искать разные источники данных. В каких-то случаях импорт был неудачным, приходили бесполезные данные. В каких-то случаях попытка импорта вовсе приводила к прекращению работы Exel. В качество интернет-сайта с данными я выбрал сайт с расписанием поездов (в этом случае получилось собрать более-менее полезные данные):

[**https://pass.rw.by/ru/tablo/?set\_exp=2100001%20**](https://pass.rw.by/ru/tablo/?set_exp=2100001%20)

****

**Копия экрана Microsoft Excel после импорта данных.**



Использованные источники:

1. <https://data.gov.ru/>

2. <http://data.worldbank.org/>

3. <https://pass.rw.by/ru/tablo/?set_exp=2100001>