

# Отчёт по ЛР №1 по ТМО

Водка Игорь, ИУ5-61

## 1) Текстовое описание набора данных

В качестве набора данных возьмём набор World University Rankings с Kaggle: <https://www.kaggle.com/mylesoneill/world-university-rankings/home> (<https://www.kaggle.com/mylesoneill/world-university-rankings/home>) Мне интересна эта тема.

Датасет состоит лишь из одного файла: `timesData.csv`, который нам предстоит разделить на обучающую и тестовую выборки. В выборке содержится множество колонок:

- `world_rank` - рейтинг (или вилка)
- `university_name` - название университета
- `country` - страна
- `teaching` - рейтинг качества преподавания
- `international` - рейтинг международного признания
- `research` - рейтинг исследований
- `citations` - industry income (knowledge transfer)
- `income` - доход
- `total_score` - общий счёт в рейтинге (целевой признак!)
- `num_students` - число студентов
- `student_staff_ratio` - кол-во студентов на одного преподавателя
- `international_students` - процент студентов из других стран
- `female_male_ratio` - отношение количества студентов мужского пола к женскому
- `year` - бесполезно

```
In [1]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
sns.set(style="ticks")
```

## Делим выборку на обучающую и тестовую

```
In [2]: from sklearn.model_selection import train_test_split

# Будем анализировать данные только на обучающей выборке
data = pd.read_csv('timesData.csv', sep=",", thousands=',')
# train, test = train_test_split(df, test_size=0.2)
```

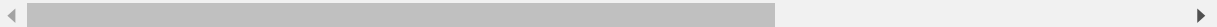
## 2) Основные характеристики датасета

Так, давайте посмотрим...

```
In [3]: # Первые 5 строк датасета
data.head()
```

```
Out[3]:
```

|   | world_rank | university_name                       | country                  | teaching | international | research | citations | income | total_score | num_students |
|---|------------|---------------------------------------|--------------------------|----------|---------------|----------|-----------|--------|-------------|--------------|
| 0 | 1          | Harvard University                    | United States of America | 99.7     | 72.4          | 98.7     | 98.8      | 34.5   | 96.1        | 20152.0      |
| 1 | 2          | California Institute of Technology    | United States of America | 97.7     | 54.6          | 98.0     | 99.9      | 83.7   | 96.0        | 2243.0       |
| 2 | 3          | Massachusetts Institute of Technology | United States of America | 97.8     | 82.3          | 91.4     | 99.9      | 87.5   | 95.6        | 11074.0      |
| 3 | 4          | Stanford University                   | United States of America | 98.3     | 29.5          | 98.1     | 99.2      | 64.3   | 94.3        | 15596.0      |
| 4 | 5          | Princeton University                  | United States of America | 90.9     | 70.3          | 95.4     | 99.9      | -      | 94.2        | 7929.0       |



```
In [4]: # Размер датасета
data.shape
```

```
Out[4]: (2603, 14)
```

```
In [5]: total_count = data.shape[0]
print('Всего строк: {}'.format(total_count))
```

```
Всего строк: 2603
```

```
In [6]: # Список колонок
data.columns
```

```
Out[6]: Index(['world_rank', 'university_name', 'country', 'teaching', 'international',
               'research', 'citations', 'income', 'total_score', 'num_students',
               'student_staff_ratio', 'international_students', 'female_male_ratio',
               'year'],
              dtype='object')
```

```
In [7]: # Проверим наличие пустых значений
# Цикл по колонкам датасета
for col in data.columns:
    # Количество пустых значений - все значения заполнены
    temp_null_count = data[data[col].isnull()].shape[0]
    print('{} - {}'.format(col, temp_null_count))
```

```
world_rank - 0
university_name - 0
country - 0
teaching - 0
international - 0
research - 0
citations - 0
income - 0
total_score - 0
num_students - 59
student_staff_ratio - 59
international_students - 67
female_male_ratio - 233
year - 0
```

```
In [8]: # Многовато... Давайте пока выкинем NaN просто.
```

```
data = data.fillna(data.mean())  
data = data.dropna()
```

```
# Немного уменьшилась наша выборка... ну и ладно.  
data.shape
```

```
Out[8]: (2362, 14)
```

```
In [9]: # Интерполируем international, research, citations, income, total_score
```

```
def adjust_total_score(df_line):  
    cols = ['international', 'research', 'citations', 'income', 'total_score']  
  
    for col in cols:  
        if df_line[col] == '-':  
            # get all non-hyphen similar rating cols  
            non_empty_cols = filter(lambda x: x != '-', map(lambda colName: df_line[colName], cols))  
            # floatify  
            non_empty_cols = list(map(lambda s: float(s), non_empty_cols))  
            # interpolate  
            df_line[col] = sum(non_empty_cols) / len(non_empty_cols)  
  
    return df_line  
  
fixed_data = data.apply(adjust_total_score, axis=1)
```

```
In [10]: fixed_data
```

Out[10]:

|    | world_rank | university_name                                   | country                  | teaching | international | research | citations | income | total_score | num_stu |
|----|------------|---------------------------------------------------|--------------------------|----------|---------------|----------|-----------|--------|-------------|---------|
| 1  | 2          | California Institute of Technology                | United States of America | 97.7     | 54.6          | 98.0     | 99.9      | 83.7   | 96.0        | 1       |
| 2  | 3          | Massachusetts Institute of Technology             | United States of America | 97.8     | 82.3          | 91.4     | 99.9      | 87.5   | 95.6        | 1       |
| 3  | 4          | Stanford University                               | United States of America | 98.3     | 29.5          | 98.1     | 99.2      | 64.3   | 94.3        | 1       |
| 4  | 5          | Princeton University                              | United States of America | 90.9     | 70.3          | 95.4     | 99.9      | 89.95  | 94.2        | 1       |
| 5  | 6          | University of Cambridge                           | United Kingdom           | 90.5     | 77.7          | 94.1     | 94.0      | 57.0   | 91.2        | 1       |
| 6  | 6          | University of Oxford                              | United Kingdom           | 88.2     | 77.2          | 93.9     | 95.1      | 73.5   | 91.2        | 1       |
| 7  | 8          | University of California, Berkeley                | United States of America | 84.2     | 39.6          | 99.3     | 97.8      | 81.95  | 91.1        | 3       |
| 8  | 9          | Imperial College London                           | United Kingdom           | 89.2     | 90.0          | 94.5     | 88.3      | 92.9   | 90.6        | 1       |
| 9  | 10         | Yale University                                   | United States of America | 92.1     | 59.2          | 89.7     | 91.5      | 82.475 | 89.5        | 1       |
| 10 | 11         | University of California, Los Angeles             | United States of America | 83.0     | 48.1          | 92.9     | 93.2      | 80.475 | 87.7        | 3       |
| 11 | 12         | University of Chicago                             | United States of America | 79.1     | 62.8          | 87.9     | 96.9      | 83.625 | 86.9        | 1       |
| 12 | 13         | Johns Hopkins University                          | United States of America | 80.9     | 58.5          | 89.2     | 92.3      | 100.0  | 86.4        | 1       |
| 13 | 14         | Cornell University                                | United States of America | 82.2     | 62.4          | 88.8     | 88.1      | 34.7   | 83.9        | 2       |
| 14 | 15         | ETH Zurich – Swiss Federal Institute of Techno... | Switzerland              | 77.5     | 93.7          | 87.8     | 83.1      | 87     | 83.4        | 1       |
| 15 | 15         | University of Michigan                            | United States of America | 83.9     | 53.3          | 89.1     | 84.1      | 59.6   | 83.4        | 4       |
| 18 | 19         | University of Pennsylvania                        | United States of America | 71.8     | 32.9          | 82.7     | 93.6      | 43.7   | 79.5        | 2       |
| 19 | 20         | Carnegie Mellon University                        | United States of America | 70.3     | 39.1          | 79.3     | 95.7      | 53.7   | 79.3        | 1       |
| 20 | 21         | University of Hong Kong                           | Hong Kong                | 68.4     | 91.4          | 71.4     | 96.1      | 56.5   | 79.2        | 1       |
| 21 | 22         | University College London                         | United Kingdom           | 74.0     | 90.8          | 81.6     | 80.6      | 39.0   | 78.4        | 2       |
| 22 | 23         | University of Washington                          | United States of America | 68.2     | 49.0          | 77.1     | 95.9      | 32.8   | 78.0        | 4       |
| 23 | 24         | Duke University                                   | United States of America | 66.8     | 49.4          | 71.5     | 92.3      | 100.0  | 76.5        | 1       |
| 24 | 25         | Northwestern University                           | United States of America | 64.5     | 60.5          | 68.8     | 95.3      | 75.125 | 75.9        | 1       |
| 26 | 27         | Georgia Institute of Technology                   | United States of America | 67.9     | 73.2          | 72.6     | 83.2      | 95.1   | 75.3        | 1       |
| 27 | 28         | Pohang University of Science and Technology       | South Korea              | 69.5     | 32.6          | 62.5     | 96.5      | 100.0  | 75.1        | 1       |

|      | world_rank | university_name                                   | country                  | teaching | international | research | citations | income  | total_score | num_stu |
|------|------------|---------------------------------------------------|--------------------------|----------|---------------|----------|-----------|---------|-------------|---------|
| 28   | 29         | University of California, Santa Barbara           | United States of America | 56.6     | 64.3          | 68.0     | 98.8      | 89.8    | 75.0        | 20      |
| 29   | 30         | University of British Columbia                    | Canada                   | 65.1     | 93.3          | 74.8     | 80.3      | 42.6    | 73.8        | 50      |
| 30   | 30         | University of North Carolina at Chapel Hill       | United States of America | 70.9     | 21.5          | 75.1     | 85.0      | 50.2    | 73.8        | 20      |
| 31   | 32         | University of California, San Diego               | United States of America | 59.8     | 31.6          | 76.3     | 90.8      | 51.8    | 73.2        | 20      |
| 32   | 33         | University of Illinois at Urbana-Champaign        | United States of America | 68.1     | 55.9          | 80.9     | 72.9      | 70.675  | 73.0        | 40      |
| 33   | 34         | National University of Singapore                  | Singapore                | 65.5     | 97.8          | 72.6     | 78.7      | 40.5    | 72.9        | 30      |
| ...  | ...        | ...                                               | ...                      | ...      | ...           | ...      | ...       | ...     | ...         | ...     |
| 2572 | 601-800    | Technical University of Madrid                    | Spain                    | 21.8     | 39.5          | 14.6     | 24.5      | 38.3    | 29.225      | 40      |
| 2573 | 601-800    | University of Tehran                              | Iran                     | 26.1     | 16.5          | 16.9     | 15.8      | 16.4    | 16.4        | 50      |
| 2574 | 601-800    | University of Texas at El Paso                    | United States of America | 18.6     | 30.4          | 18.7     | 18.4      | 22.5    | 22.5        | 10      |
| 2575 | 601-800    | Texas Tech University                             | United States of America | 27.9     | 36.8          | 17.2     | 22.0      | 25.3333 | 25.3333     | 20      |
| 2576 | 601-800    | Tokai University                                  | Japan                    | 17.9     | 19.3          | 7.6      | 15.3      | 34.4    | 19.15       | 20      |
| 2577 | 601-800    | Tokushima University                              | Japan                    | 25.3     | 16.8          | 21.6     | 12.8      | 59.6    | 27.7        | 10      |
| 2578 | 601-800    | Tokyo University of Marine Science and Technology | Japan                    | 27.9     | 24.5          | 12.4     | 7.7       | 57.9    | 25.625      | 10      |
| 2579 | 601-800    | Tokyo University of Science                       | Japan                    | 23.0     | 15.4          | 24.1     | 21.4      | 37.6    | 24.625      | 20      |
| 2580 | 601-800    | Tomsk State University                            | Russian Federation       | 34.8     | 36.9          | 20.8     | 7.6       | 44.0    | 27.325      | 10      |
| 2581 | 601-800    | Tottori University                                | Japan                    | 24.3     | 16.7          | 10.1     | 9.6       | 34.5    | 17.725      | 10      |
| 2582 | 601-800    | Toyohashi University of Technology                | Japan                    | 22.0     | 25.4          | 18.9     | 15.8      | 50.3    | 27.6        | 10      |
| 2583 | 601-800    | Universiti Kebangsaan Malaysia                    | Malaysia                 | 24.3     | 29.7          | 15.9     | 10.9      | 28.4    | 21.225      | 20      |
| 2584 | 601-800    | Universiti Putra Malaysia                         | Malaysia                 | 25.3     | 50.1          | 20.9     | 10.2      | 34.2    | 28.85       | 20      |
| 2585 | 601-800    | Universiti Sains Malaysia                         | Malaysia                 | 26.9     | 44.2          | 16.6     | 12.4      | 34.4    | 26.9        | 20      |
| 2586 | 601-800    | Universiti Teknologi MARA                         | Malaysia                 | 15.2     | 14.8          | 7.7      | 18.2      | 28.3    | 17.25       | 60      |
| 2587 | 601-800    | Ural Federal University                           | Russian Federation       | 24.8     | 17.3          | 10.6     | 16.8      | 35.6    | 20.075      | 20      |
| 2588 | 601-800    | V.N. Karazin Kharkiv National University          | Ukraine                  | 21.7     | 48.4          | 8.9      | 1.7       | 28.8    | 21.95       | 10      |
| 2589 | 601-800    | University of Vigo                                | Spain                    | 18.4     | 30.7          | 10.5     | 31.8      | 38.1    | 27.775      | 20      |
| 2590 | 601-800    | Vilnius University                                | Lithuania                | 18.3     | 40.8          | 13.6     | 26.1      | 41.0    | 30.375      | 10      |
| 2591 | 601-800    | Warsaw University of Technology                   | Poland                   | 19.4     | 20.7          | 8.5      | 40.3      | 47.4    | 29.225      | 30      |
| 2592 | 601-800    | Waseda University                                 | Japan                    | 23.6     | 29.7          | 14.6     | 29.4      | 32.4    | 26.525      | 50      |

|      | world_rank | university_name                   | country        | teaching | international | research | citations | income  | total_score | num_st |
|------|------------|-----------------------------------|----------------|----------|---------------|----------|-----------|---------|-------------|--------|
| 2593 | 601-800    | University of West Bohemia        | Czech Republic | 16.3     | 23.1          | 9.7      | 29.8      | 32.1    | 23.675      | 1      |
| 2594 | 601-800    | University of the West of England | United Kingdom | 16.9     | 48.5          | 11.2     | 34.6      | 28.5    | 30.7        | 2      |
| 2595 | 601-800    | West University of Timișoara      | Romania        | 16.1     | 21.0          | 3.9      | 22.4      | 15.7667 | 15.7667     | 1      |
| 2596 | 601-800    | University of Westminster         | United Kingdom | 17.3     | 81.9          | 11.7     | 21.1      | 28.5    | 35.8        | 1      |
| 2597 | 601-800    | Xidian University                 | China          | 17.9     | 12.8          | 12.1     | 8.9       | 83.7    | 29.375      | 3      |
| 2598 | 601-800    | Yeungnam University               | South Korea    | 18.6     | 24.3          | 10.9     | 26.5      | 35.4    | 24.275      | 2      |
| 2599 | 601-800    | Yildiz Technical University       | Turkey         | 14.5     | 14.9          | 7.6      | 19.3      | 44.0    | 21.45       | 3      |
| 2601 | 601-800    | Yokohama National University      | Japan          | 20.1     | 23.3          | 16.0     | 13.5      | 40.4    | 23.3        | 1      |
| 2602 | 601-800    | Yuan Ze University                | Taiwan         | 16.2     | 17.7          | 18.3     | 28.6      | 39.8    | 26.1        | 1      |

2362 rows × 14 columns



```
In [11]: # Основные статистические характеристики набора данных
fixed_data.describe()
```

Out[11]:

|       | teaching    | research    | citations   | num_students  | student_staff_ratio | year        |
|-------|-------------|-------------|-------------|---------------|---------------------|-------------|
| count | 2362.000000 | 2362.000000 | 2362.000000 | 2362.000000   | 2362.000000         | 2362.000000 |
| mean  | 37.146190   | 35.310288   | 61.004953   | 23845.077053  | 18.707282           | 2014.092295 |
| std   | 17.145579   | 20.876934   | 23.091455   | 18008.833624  | 11.530523           | 1.682795    |
| min   | 9.900000    | 2.900000    | 1.200000    | 462.000000    | 0.600000            | 2011.000000 |
| 25%   | 24.500000   | 19.400000   | 45.500000   | 12551.000000  | 12.200000           | 2013.000000 |
| 50%   | 33.100000   | 30.100000   | 62.700000   | 20584.000000  | 16.300000           | 2014.000000 |
| 75%   | 45.700000   | 46.200000   | 79.200000   | 30333.000000  | 21.900000           | 2016.000000 |
| max   | 98.300000   | 99.400000   | 100.000000  | 379231.000000 | 162.600000          | 2016.000000 |

```
In [12]: # Определим уникальные значения для целевого признака
data['world_rank'].unique()
```

Out[12]: array(['2', '3', '4', '5', '6', '8', '9', '10', '11', '12', '13', '14', '15', '19', '20', '21', '22', '23', '24', '25', '27', '28', '29', '30', '32', '33', '34', '35', '36', '39', '40', '42', '43', '47', '48', '49', '51', '52', '53', '54', '56', '58', '59', '60', '61', '63', '64', '65', '66', '67', '68', '71', '72', '73', '75', '76', '77', '78', '79', '81', '83', '85', '87', '88', '90', '93', '95', '98', '99', '100', '101', '102', '103', '104', '105', '106', '107', '109', '111', '112', '114', '115', '117', '118', '119', '120', '122', '124', '127', '128', '129', '130', '132', '135', '137', '138', '139', '140', '142', '143', '144', '145', '147', '149', '151', '152', '155', '156', '159', '161', '163', '164', '165', '167', '168', '170', '171', '172', '173', '174', '177', '178', '181', '182', '183', '184', '185', '186', '187', '189', '190', '193', '195', '196', '197', '199', '1', '7', '16', '17', '18', '26', '31', '37', '38', '44', '45', '46', '55', '57', '69', '70', '74', '84', '86', '89', '91', '92', '94', '96', '97', '108', '110', '113', '121', '123', '125', '131', '133', '134', '141', '146', '148', '150', '154', '157', '162', '169', '176', '180', '191', '194', '200', '201-225', '226-250', '251-275', '276-300', '301-350', '350-400', '41', '50', '62', '80', '116', '153', '158', '166', '192', '198', '351-400', '126', '136', '160', '188', '82', '175', '=39', '=44', '=47', '=56', '=60', '=65', '=76', '=82', '=88', '=90', '=94', '=99', '=101', '=104', '=106', '=110', '=113', '=120', '=123', '=125', '=127', '=131', '=133', '=138', '=144', '=149', '=158', '=161', '=164', '=167', '=172', '=176', '179', '=180', '=182', '=185', '=190', '=193', '=196', '201-250', '251-300', '401-500', '501-600', '601-800'], dtype=object)

```
In [13]: fixed_data
```



Out[13]:

|    | world_rank | university_name                                   | country                  | teaching | international | research | citations | income | total_score | num_stu |
|----|------------|---------------------------------------------------|--------------------------|----------|---------------|----------|-----------|--------|-------------|---------|
| 1  | 2          | California Institute of Technology                | United States of America | 97.7     | 54.6          | 98.0     | 99.9      | 83.7   | 96.0        | 1       |
| 2  | 3          | Massachusetts Institute of Technology             | United States of America | 97.8     | 82.3          | 91.4     | 99.9      | 87.5   | 95.6        | 1       |
| 3  | 4          | Stanford University                               | United States of America | 98.3     | 29.5          | 98.1     | 99.2      | 64.3   | 94.3        | 1       |
| 4  | 5          | Princeton University                              | United States of America | 90.9     | 70.3          | 95.4     | 99.9      | 89.95  | 94.2        | 1       |
| 5  | 6          | University of Cambridge                           | United Kingdom           | 90.5     | 77.7          | 94.1     | 94.0      | 57.0   | 91.2        | 1       |
| 6  | 6          | University of Oxford                              | United Kingdom           | 88.2     | 77.2          | 93.9     | 95.1      | 73.5   | 91.2        | 1       |
| 7  | 8          | University of California, Berkeley                | United States of America | 84.2     | 39.6          | 99.3     | 97.8      | 81.95  | 91.1        | 3       |
| 8  | 9          | Imperial College London                           | United Kingdom           | 89.2     | 90.0          | 94.5     | 88.3      | 92.9   | 90.6        | 1       |
| 9  | 10         | Yale University                                   | United States of America | 92.1     | 59.2          | 89.7     | 91.5      | 82.475 | 89.5        | 1       |
| 10 | 11         | University of California, Los Angeles             | United States of America | 83.0     | 48.1          | 92.9     | 93.2      | 80.475 | 87.7        | 3       |
| 11 | 12         | University of Chicago                             | United States of America | 79.1     | 62.8          | 87.9     | 96.9      | 83.625 | 86.9        | 1       |
| 12 | 13         | Johns Hopkins University                          | United States of America | 80.9     | 58.5          | 89.2     | 92.3      | 100.0  | 86.4        | 1       |
| 13 | 14         | Cornell University                                | United States of America | 82.2     | 62.4          | 88.8     | 88.1      | 34.7   | 83.9        | 2       |
| 14 | 15         | ETH Zurich – Swiss Federal Institute of Techno... | Switzerland              | 77.5     | 93.7          | 87.8     | 83.1      | 87     | 83.4        | 1       |
| 15 | 15         | University of Michigan                            | United States of America | 83.9     | 53.3          | 89.1     | 84.1      | 59.6   | 83.4        | 4       |
| 18 | 19         | University of Pennsylvania                        | United States of America | 71.8     | 32.9          | 82.7     | 93.6      | 43.7   | 79.5        | 2       |
| 19 | 20         | Carnegie Mellon University                        | United States of America | 70.3     | 39.1          | 79.3     | 95.7      | 53.7   | 79.3        | 1       |
| 20 | 21         | University of Hong Kong                           | Hong Kong                | 68.4     | 91.4          | 71.4     | 96.1      | 56.5   | 79.2        | 1       |
| 21 | 22         | University College London                         | United Kingdom           | 74.0     | 90.8          | 81.6     | 80.6      | 39.0   | 78.4        | 2       |
| 22 | 23         | University of Washington                          | United States of America | 68.2     | 49.0          | 77.1     | 95.9      | 32.8   | 78.0        | 4       |
| 23 | 24         | Duke University                                   | United States of America | 66.8     | 49.4          | 71.5     | 92.3      | 100.0  | 76.5        | 1       |
| 24 | 25         | Northwestern University                           | United States of America | 64.5     | 60.5          | 68.8     | 95.3      | 75.125 | 75.9        | 1       |
| 26 | 27         | Georgia Institute of Technology                   | United States of America | 67.9     | 73.2          | 72.6     | 83.2      | 95.1   | 75.3        | 1       |
| 27 | 28         | Pohang University of Science and Technology       | South Korea              | 69.5     | 32.6          | 62.5     | 96.5      | 100.0  | 75.1        | 1       |

|      | world_rank | university_name                                   | country                  | teaching | international | research | citations | income  | total_score | num_stu |
|------|------------|---------------------------------------------------|--------------------------|----------|---------------|----------|-----------|---------|-------------|---------|
| 28   | 29         | University of California, Santa Barbara           | United States of America | 56.6     | 64.3          | 68.0     | 98.8      | 89.8    | 75.0        | 20      |
| 29   | 30         | University of British Columbia                    | Canada                   | 65.1     | 93.3          | 74.8     | 80.3      | 42.6    | 73.8        | 50      |
| 30   | 30         | University of North Carolina at Chapel Hill       | United States of America | 70.9     | 21.5          | 75.1     | 85.0      | 50.2    | 73.8        | 20      |
| 31   | 32         | University of California, San Diego               | United States of America | 59.8     | 31.6          | 76.3     | 90.8      | 51.8    | 73.2        | 20      |
| 32   | 33         | University of Illinois at Urbana-Champaign        | United States of America | 68.1     | 55.9          | 80.9     | 72.9      | 70.675  | 73.0        | 40      |
| 33   | 34         | National University of Singapore                  | Singapore                | 65.5     | 97.8          | 72.6     | 78.7      | 40.5    | 72.9        | 30      |
| ...  | ...        | ...                                               | ...                      | ...      | ...           | ...      | ...       | ...     | ...         | ...     |
| 2572 | 601-800    | Technical University of Madrid                    | Spain                    | 21.8     | 39.5          | 14.6     | 24.5      | 38.3    | 29.225      | 40      |
| 2573 | 601-800    | University of Tehran                              | Iran                     | 26.1     | 16.5          | 16.9     | 15.8      | 16.4    | 16.4        | 50      |
| 2574 | 601-800    | University of Texas at El Paso                    | United States of America | 18.6     | 30.4          | 18.7     | 18.4      | 22.5    | 22.5        | 10      |
| 2575 | 601-800    | Texas Tech University                             | United States of America | 27.9     | 36.8          | 17.2     | 22.0      | 25.3333 | 25.3333     | 20      |
| 2576 | 601-800    | Tokai University                                  | Japan                    | 17.9     | 19.3          | 7.6      | 15.3      | 34.4    | 19.15       | 20      |
| 2577 | 601-800    | Tokushima University                              | Japan                    | 25.3     | 16.8          | 21.6     | 12.8      | 59.6    | 27.7        | 10      |
| 2578 | 601-800    | Tokyo University of Marine Science and Technology | Japan                    | 27.9     | 24.5          | 12.4     | 7.7       | 57.9    | 25.625      | 10      |
| 2579 | 601-800    | Tokyo University of Science                       | Japan                    | 23.0     | 15.4          | 24.1     | 21.4      | 37.6    | 24.625      | 20      |
| 2580 | 601-800    | Tomsk State University                            | Russian Federation       | 34.8     | 36.9          | 20.8     | 7.6       | 44.0    | 27.325      | 10      |
| 2581 | 601-800    | Tottori University                                | Japan                    | 24.3     | 16.7          | 10.1     | 9.6       | 34.5    | 17.725      | 10      |
| 2582 | 601-800    | Toyohashi University of Technology                | Japan                    | 22.0     | 25.4          | 18.9     | 15.8      | 50.3    | 27.6        | 10      |
| 2583 | 601-800    | Universiti Kebangsaan Malaysia                    | Malaysia                 | 24.3     | 29.7          | 15.9     | 10.9      | 28.4    | 21.225      | 20      |
| 2584 | 601-800    | Universiti Putra Malaysia                         | Malaysia                 | 25.3     | 50.1          | 20.9     | 10.2      | 34.2    | 28.85       | 20      |
| 2585 | 601-800    | Universiti Sains Malaysia                         | Malaysia                 | 26.9     | 44.2          | 16.6     | 12.4      | 34.4    | 26.9        | 20      |
| 2586 | 601-800    | Universiti Teknologi MARA                         | Malaysia                 | 15.2     | 14.8          | 7.7      | 18.2      | 28.3    | 17.25       | 60      |
| 2587 | 601-800    | Ural Federal University                           | Russian Federation       | 24.8     | 17.3          | 10.6     | 16.8      | 35.6    | 20.075      | 20      |
| 2588 | 601-800    | V.N. Karazin Kharkiv National University          | Ukraine                  | 21.7     | 48.4          | 8.9      | 1.7       | 28.8    | 21.95       | 10      |
| 2589 | 601-800    | University of Vigo                                | Spain                    | 18.4     | 30.7          | 10.5     | 31.8      | 38.1    | 27.775      | 20      |
| 2590 | 601-800    | Vilnius University                                | Lithuania                | 18.3     | 40.8          | 13.6     | 26.1      | 41.0    | 30.375      | 10      |
| 2591 | 601-800    | Warsaw University of Technology                   | Poland                   | 19.4     | 20.7          | 8.5      | 40.3      | 47.4    | 29.225      | 30      |
| 2592 | 601-800    | Waseda University                                 | Japan                    | 23.6     | 29.7          | 14.6     | 29.4      | 32.4    | 26.525      | 50      |

|      | world_rank | university_name                   | country        | teaching | international | research | citations | income  | total_score | num_stu |
|------|------------|-----------------------------------|----------------|----------|---------------|----------|-----------|---------|-------------|---------|
| 2593 | 601-800    | University of West Bohemia        | Czech Republic | 16.3     | 23.1          | 9.7      | 29.8      | 32.1    | 23.675      | 1       |
| 2594 | 601-800    | University of the West of England | United Kingdom | 16.9     | 48.5          | 11.2     | 34.6      | 28.5    | 30.7        | 2       |
| 2595 | 601-800    | West University of Timișoara      | Romania        | 16.1     | 21.0          | 3.9      | 22.4      | 15.7667 | 15.7667     | 1       |
| 2596 | 601-800    | University of Westminster         | United Kingdom | 17.3     | 81.9          | 11.7     | 21.1      | 28.5    | 35.8        | 1       |
| 2597 | 601-800    | Xidian University                 | China          | 17.9     | 12.8          | 12.1     | 8.9       | 83.7    | 29.375      | 3       |
| 2598 | 601-800    | Yeungnam University               | South Korea    | 18.6     | 24.3          | 10.9     | 26.5      | 35.4    | 24.275      | 2       |
| 2599 | 601-800    | Yildiz Technical University       | Turkey         | 14.5     | 14.9          | 7.6      | 19.3      | 44.0    | 21.45       | 3       |
| 2601 | 601-800    | Yokohama National University      | Japan          | 20.1     | 23.3          | 16.0     | 13.5      | 40.4    | 23.3        | 1       |
| 2602 | 601-800    | Yuan Ze University                | Taiwan         | 16.2     | 17.7          | 18.3     | 28.6      | 39.8    | 26.1        | 1       |

2362 rows × 14 columns



In [14]: *# Чуть поправляем данные*

```
fixed_data.international_students = fixed_data.international_students.str.replace('%',
'')

numeric_cols = ['teaching', 'international', 'research', 'citations', 'income', 'total_s
core', 'num_students', 'international_students']
fixed_data[numeric_cols] = fixed_data[numeric_cols].apply(pd.to_numeric)

fixed_data.dtypes

def fix_female_male_ratio(x):
    if x == '-':
        return 1
    else:
        split = x.split(' : ')
        f = int(split[0])
        m = int(split[1])
        if m == 0:
            m = 0.01
        return f / m

def fix_world_rank(x):
    return int(x.split('-')[0].replace('=', ''))

fixed_data.female_male_ratio = fixed_data.female_male_ratio.apply(fix_female_male_ratio)
fixed_data.world_rank = fixed_data.world_rank.apply(fix_world_rank)
```

```
In [15]: def group_all_years(years):
    sorted_years = years.sort_values('year', ascending=False)
    first_year = sorted_years.head(1)
    last_year = sorted_years.tail(1)
    for col in ['teaching', 'international', 'research', 'citations', 'income', 'total_s
core', 'num_students', 'international_students']:
        last_year[col + '_progress'] = years.iloc[0][col] / years.iloc[-1][col]
    return last_year

fixed_data_with_progress = fixed_data.groupby('university_name').apply(group_all_years)
fixed_data_with_progress = fixed_data_with_progress.drop('year', axis=1)
fixed_data_with_progress
```

```
/usr/local/lib/python3.6/dist-packages/ipykernel_launcher.py:6: RuntimeWarning: invalid v  
alue encountered in long_scalars
```

Out[15]:

|                                             |      | world_rank | university_name                             | country                  | teaching | international | research | citations | income     |
|---------------------------------------------|------|------------|---------------------------------------------|--------------------------|----------|---------------|----------|-----------|------------|
| university_name                             |      |            |                                             |                          |          |               |          |           |            |
| AGH University of Science and Technology    | 2405 | 601        | AGH University of Science and Technology    | Poland                   | 14.2     | 17.9          | 3.7      | 35.7      | 19.100000  |
| Aalborg University                          | 501  | 301        | Aalborg University                          | Denmark                  | 19.0     | 75.3          | 20.0     | 27.1      | 36.400000  |
| Aalto University                            | 502  | 301        | Aalto University                            | Finland                  | 26.2     | 49.0          | 22.2     | 37.5      | 61.900000  |
| Aarhus University                           | 166  | 167        | Aarhus University                           | Denmark                  | 38.1     | 33.4          | 55.6     | 57.3      | 61.500000  |
| Aberystwyth University                      | 476  | 276        | Aberystwyth University                      | United Kingdom           | 19.8     | 63.8          | 15.5     | 56.6      | 35.500000  |
| Adam Mickiewicz University                  | 2404 | 601        | Adam Mickiewicz University                  | Poland                   | 20.0     | 25.7          | 11.0     | 15.3      | 28.700000  |
| Aix-Marseille University                    | 2057 | 251        | Aix-Marseille University                    | France                   | 36.7     | 63.0          | 22.1     | 64.9      | 33.100000  |
| Ajou University                             | 2406 | 601        | Ajou University                             | South Korea              | 19.5     | 20.0          | 11.9     | 23.9      | 45.700000  |
| Alexandria University                       | 146  | 147        | Alexandria University                       | Egypt                    | 29.5     | 19.3          | 28.0     | 99.8      | 36.000000  |
| Alexandru Ioan Cuza University              | 2409 | 601        | Alexandru Ioan Cuza University              | Romania                  | 24.9     | 46.9          | 13.6     | 7.0       | 28.200000  |
| Aligarh Muslim University                   | 2410 | 601        | Aligarh Muslim University                   | India                    | 28.3     | 18.7          | 10.0     | 20.9      | 29.600000  |
| American University                         | 2204 | 401        | American University                         | United States of America | 42.2     | 28.9          | 16.5     | 41.1      | 35.900000  |
| American University of Beirut               | 2303 | 501        | American University of Beirut               | Lebanon                  | 27.7     | 93.0          | 11.2     | 31.9      | 45.366667  |
| American University of Sharjah              | 2411 | 601        | American University of Sharjah              | United Arab Emirates     | 12.4     | 95.6          | 10.6     | 13.3      | 33.300000  |
| Amirkabir University of Technology          | 2304 | 501        | Amirkabir University of Technology          | Iran                     | 24.5     | 7.7           | 25.7     | 34.8      | 55.700000  |
| Anadolu University                          | 2413 | 601        | Anadolu University                          | Turkey                   | 12.2     | 14.3          | 22.6     | 10.9      | 100.000000 |
| Andhra University                           | 2414 | 601        | Andhra University                           | India                    | 34.8     | 7.2           | 6.9      | 1.2       | 31.300000  |
| Aristotle University of Thessaloniki        | 2416 | 601        | Aristotle University of Thessaloniki        | Greece                   | 22.6     | 36.6          | 15.0     | 29.5      | 33.600000  |
| Arizona State University                    | 160  | 161        | Arizona State University                    | United States of America | 43.0     | 24.1          | 44.1     | 66.9      | 46.350000  |
| Asia University, Taiwan                     | 2417 | 601        | Asia University, Taiwan                     | Taiwan                   | 14.4     | 18.3          | 15.9     | 30.5      | 38.400000  |
| Athens University of Economics and Business | 2418 | 601        | Athens University of Economics and Business | Greece                   | 13.4     | 39.2          | 18.8     | 31.2      | 66.500000  |
| Auburn University                           | 552  | 350        | Auburn University                           | United States of America | 33.7     | 22.5          | 18.7     | 10.3      | 47.300000  |
| Auckland University of Technology           | 2419 | 601        | Auckland University of Technology           | New Zealand              | 17.3     | 95.6          | 9.8      | 21.7      | 28.500000  |
| Austral University of Chile                 | 2420 | 601        | Austral University of Chile                 | Chile                    | 16.7     | 45.2          | 9.7      | 14.2      | 41.100000  |
| Australian National University              | 42   | 43         | Australian National University              | Australia                | 51.9     | 93.9          | 62.4     | 81.0      | 76.075000  |
| Autonomous University of Barcelona          | 400  | 201        | Autonomous University of Barcelona          | Spain                    | 33.7     | 45.9          | 27.9     | 57.9      | 37.000000  |

|                                                     |      | world_rank | university_name                                   | country                  | teaching | international | research | citations | income    |
|-----------------------------------------------------|------|------------|---------------------------------------------------|--------------------------|----------|---------------|----------|-----------|-----------|
| university_name                                     |      |            |                                                   |                          |          |               |          |           |           |
| Autonomous University of Madrid                     | 477  | 276        | Autonomous University of Madrid                   | Spain                    | 28.8     | 39.7          | 21.4     | 47.5      | 32.500000 |
| Babeş-Bolyai University                             | 2308 | 501        | Babeş-Bolyai University                           | Romania                  | 27.9     | 35.4          | 12.5     | 32.1      | 28.600000 |
| Bangor University                                   | 450  | 251        | Bangor University                                 | United Kingdom           | 24.9     | 67.1          | 23.8     | 48.0      | 29.000000 |
| Bar-Ilan University                                 | 504  | 301        | Bar-Ilan University                               | Israel                   | 27.9     | 47.0          | 32.2     | 15.7      | 28.500000 |
| ...                                                 | ...  | ...        | ...                                               | ...                      | ...      | ...           | ...      | ...       | ...       |
| Vilnius University                                  | 2590 | 601        | Vilnius University                                | Lithuania                | 18.3     | 40.8          | 13.6     | 26.1      | 41.000000 |
| Virginia Polytechnic Institute and State University | 474  | 251        | Virginia Polytechnic Institute and State Unive... | United States of America | 36.9     | 25.1          | 40.4     | 30.3      | 24.200000 |
| Vrije Universiteit Brussel                          | 548  | 301        | Vrije Universiteit Brussel                        | Belgium                  | 23.4     | 49.4          | 21.1     | 36.7      | 66.200000 |
| VŠB - Technical University of Ostrava               | 2150 | 301        | VŠB - Technical University of Ostrava             | Czech Republic           | 18.8     | 22.3          | 15.7     | 85.5      | 32.100000 |
| Wageningen University and Research Center           | 143  | 144        | Wageningen University and Research Center         | Netherlands              | 58.5     | 24.3          | 48.8     | 53.0      | 44.500000 |
| Wake Forest University                              | 91   | 90         | Wake Forest University                            | United States of America | 54.6     | 24.4          | 42.9     | 79.2      | 51.050000 |
| Warsaw University of Technology                     | 2591 | 601        | Warsaw University of Technology                   | Poland                   | 19.4     | 20.7          | 8.5      | 40.3      | 47.400000 |
| Waseda University                                   | 599  | 350        | Waseda University                                 | Japan                    | 25.4     | 27.1          | 17.3     | 29.7      | 27.300000 |
| Washington State University                         | 549  | 301        | Washington State University                       | United States of America | 28.8     | 31.0          | 24.5     | 33.5      | 34.300000 |
| Wayne State University                              | 475  | 251        | Wayne State University                            | United States of America | 34.7     | 24.6          | 16.4     | 55.2      | 32.066667 |
| West University of Timișoara                        | 2595 | 601        | West University of Timișoara                      | Romania                  | 16.1     | 21.0          | 3.9      | 22.4      | 15.766667 |
| Western Sydney University                           | 1800 | 351        | Western Sydney University                         | Australia                | 17.8     | 50.1          | 22.4     | 58.7      | 30.400000 |
| William & Mary                                      | 74   | 75         | William & Mary                                    | United States of America | 53.1     | 20.9          | 36.1     | 95.6      | 53.250000 |
| Wuhan University of Technology                      | 1353 | 301        | Wuhan University of Technology                    | China                    | 14.8     | 18.9          | 7.8      | 78.1      | 58.700000 |
| Xiamen University                                   | 2302 | 401        | Xiamen University                                 | China                    | 26.7     | 25.0          | 15.6     | 47.0      | 29.200000 |
| Xidian University                                   | 2597 | 601        | Xidian University                                 | China                    | 17.9     | 12.8          | 12.1     | 8.9       | 83.700000 |
| Xi'an Jiaotong University                           | 2401 | 501        | Xi'an Jiaotong University                         | China                    | 28.7     | 25.8          | 22.5     | 25.5      | 70.400000 |
| Yale University                                     | 9    | 10         | Yale University                                   | United States of America | 92.1     | 59.2          | 89.7     | 91.5      | 82.475000 |
| Yeshiva University                                  | 69   | 68         | Yeshiva University                                | United States of America | 63.5     | 53.3          | 46.7     | 74.4      | 58.950000 |
| Yeungnam University                                 | 2598 | 601        | Yeungnam University                               | South Korea              | 18.6     | 24.3          | 10.9     | 26.5      | 35.400000 |
| Yokohama National University                        | 2601 | 601        | Yokohama National University                      | Japan                    | 20.1     | 23.3          | 16.0     | 13.5      | 40.400000 |

|                                          | world_rank |     | university_name                          | country     | teaching | international | research | citations | income    |
|------------------------------------------|------------|-----|------------------------------------------|-------------|----------|---------------|----------|-----------|-----------|
| university_name                          |            |     |                                          |             |          |               |          |           |           |
| York University                          | 500        | 276 | York University                          | Canada      | 19.9     | 57.7          | 27.7     | 41.2      | 41.700000 |
| Yuan Ze University                       | 601        | 350 | Yuan Ze University                       | Taiwan      | 10.8     | 12.8          | 9.6      | 58.3      | 29.200000 |
| Yıldız Technical University              | 2599       | 601 | Yıldız Technical University              | Turkey      | 14.5     | 14.9          | 7.6      | 19.3      | 44.000000 |
| Zhejiang University                      | 197        | 197 | Zhejiang University                      | China       | 54.6     | 29.6          | 41.3     | 44.3      | 70.300000 |
| École Normale Supérieure                 | 41         | 42  | École Normale Supérieure                 | France      | 66.8     | 44.9          | 48.2     | 95.7      | 30.700000 |
| École Normale Supérieure de Lyon         | 99         | 100 | École Normale Supérieure de Lyon         | France      | 51.1     | 37.6          | 34.4     | 88.8      | 26.100000 |
| École Polytechnique                      | 38         | 39  | École Polytechnique                      | France      | 57.9     | 77.9          | 56.1     | 91.4      | 73.725000 |
| École Polytechnique Fédérale de Lausanne | 47         | 48  | École Polytechnique Fédérale de Lausanne | Switzerland | 55.0     | 100.0         | 56.1     | 83.8      | 38.000000 |
| Örebro University                        | 2134       | 301 | Örebro University                        | Sweden      | 18.3     | 39.4          | 10.4     | 87.8      | 29.800000 |

739 rows × 21 columns

### 3) Визуальное исследование датасета



```
In [16]: data = fixed_data_with_progress  
data
```

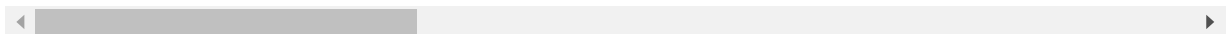
Out[16]:

|                                             |      | world_rank | university_name                             | country                  | teaching | international | research | citations | income     |
|---------------------------------------------|------|------------|---------------------------------------------|--------------------------|----------|---------------|----------|-----------|------------|
| university_name                             |      |            |                                             |                          |          |               |          |           |            |
| AGH University of Science and Technology    | 2405 | 601        | AGH University of Science and Technology    | Poland                   | 14.2     | 17.9          | 3.7      | 35.7      | 19.100000  |
| Aalborg University                          | 501  | 301        | Aalborg University                          | Denmark                  | 19.0     | 75.3          | 20.0     | 27.1      | 36.400000  |
| Aalto University                            | 502  | 301        | Aalto University                            | Finland                  | 26.2     | 49.0          | 22.2     | 37.5      | 61.900000  |
| Aarhus University                           | 166  | 167        | Aarhus University                           | Denmark                  | 38.1     | 33.4          | 55.6     | 57.3      | 61.500000  |
| Aberystwyth University                      | 476  | 276        | Aberystwyth University                      | United Kingdom           | 19.8     | 63.8          | 15.5     | 56.6      | 35.500000  |
| Adam Mickiewicz University                  | 2404 | 601        | Adam Mickiewicz University                  | Poland                   | 20.0     | 25.7          | 11.0     | 15.3      | 28.700000  |
| Aix-Marseille University                    | 2057 | 251        | Aix-Marseille University                    | France                   | 36.7     | 63.0          | 22.1     | 64.9      | 33.100000  |
| Ajou University                             | 2406 | 601        | Ajou University                             | South Korea              | 19.5     | 20.0          | 11.9     | 23.9      | 45.700000  |
| Alexandria University                       | 146  | 147        | Alexandria University                       | Egypt                    | 29.5     | 19.3          | 28.0     | 99.8      | 36.000000  |
| Alexandru Ioan Cuza University              | 2409 | 601        | Alexandru Ioan Cuza University              | Romania                  | 24.9     | 46.9          | 13.6     | 7.0       | 28.200000  |
| Aligarh Muslim University                   | 2410 | 601        | Aligarh Muslim University                   | India                    | 28.3     | 18.7          | 10.0     | 20.9      | 29.600000  |
| American University                         | 2204 | 401        | American University                         | United States of America | 42.2     | 28.9          | 16.5     | 41.1      | 35.900000  |
| American University of Beirut               | 2303 | 501        | American University of Beirut               | Lebanon                  | 27.7     | 93.0          | 11.2     | 31.9      | 45.366667  |
| American University of Sharjah              | 2411 | 601        | American University of Sharjah              | United Arab Emirates     | 12.4     | 95.6          | 10.6     | 13.3      | 33.300000  |
| Amirkabir University of Technology          | 2304 | 501        | Amirkabir University of Technology          | Iran                     | 24.5     | 7.7           | 25.7     | 34.8      | 55.700000  |
| Anadolu University                          | 2413 | 601        | Anadolu University                          | Turkey                   | 12.2     | 14.3          | 22.6     | 10.9      | 100.000000 |
| Andhra University                           | 2414 | 601        | Andhra University                           | India                    | 34.8     | 7.2           | 6.9      | 1.2       | 31.300000  |
| Aristotle University of Thessaloniki        | 2416 | 601        | Aristotle University of Thessaloniki        | Greece                   | 22.6     | 36.6          | 15.0     | 29.5      | 33.600000  |
| Arizona State University                    | 160  | 161        | Arizona State University                    | United States of America | 43.0     | 24.1          | 44.1     | 66.9      | 46.350000  |
| Asia University, Taiwan                     | 2417 | 601        | Asia University, Taiwan                     | Taiwan                   | 14.4     | 18.3          | 15.9     | 30.5      | 38.400000  |
| Athens University of Economics and Business | 2418 | 601        | Athens University of Economics and Business | Greece                   | 13.4     | 39.2          | 18.8     | 31.2      | 66.500000  |
| Auburn University                           | 552  | 350        | Auburn University                           | United States of America | 33.7     | 22.5          | 18.7     | 10.3      | 47.300000  |
| Auckland University of Technology           | 2419 | 601        | Auckland University of Technology           | New Zealand              | 17.3     | 95.6          | 9.8      | 21.7      | 28.500000  |
| Austral University of Chile                 | 2420 | 601        | Austral University of Chile                 | Chile                    | 16.7     | 45.2          | 9.7      | 14.2      | 41.100000  |
| Australian National University              | 42   | 43         | Australian National University              | Australia                | 51.9     | 93.9          | 62.4     | 81.0      | 76.075000  |
| Autonomous University of Barcelona          | 400  | 201        | Autonomous University of Barcelona          | Spain                    | 33.7     | 45.9          | 27.9     | 57.9      | 37.000000  |

|                                                     |      | world_rank | university_name                                   | country                  | teaching | international | research | citations | income    |
|-----------------------------------------------------|------|------------|---------------------------------------------------|--------------------------|----------|---------------|----------|-----------|-----------|
| university_name                                     |      |            |                                                   |                          |          |               |          |           |           |
| Autonomous University of Madrid                     | 477  | 276        | Autonomous University of Madrid                   | Spain                    | 28.8     | 39.7          | 21.4     | 47.5      | 32.500000 |
| Babeş-Bolyai University                             | 2308 | 501        | Babeş-Bolyai University                           | Romania                  | 27.9     | 35.4          | 12.5     | 32.1      | 28.600000 |
| Bangor University                                   | 450  | 251        | Bangor University                                 | United Kingdom           | 24.9     | 67.1          | 23.8     | 48.0      | 29.000000 |
| Bar-Ilan University                                 | 504  | 301        | Bar-Ilan University                               | Israel                   | 27.9     | 47.0          | 32.2     | 15.7      | 28.500000 |
| ...                                                 | ...  | ...        | ...                                               | ...                      | ...      | ...           | ...      | ...       | ...       |
| Vilnius University                                  | 2590 | 601        | Vilnius University                                | Lithuania                | 18.3     | 40.8          | 13.6     | 26.1      | 41.000000 |
| Virginia Polytechnic Institute and State University | 474  | 251        | Virginia Polytechnic Institute and State Unive... | United States of America | 36.9     | 25.1          | 40.4     | 30.3      | 24.200000 |
| Vrije Universiteit Brussel                          | 548  | 301        | Vrije Universiteit Brussel                        | Belgium                  | 23.4     | 49.4          | 21.1     | 36.7      | 66.200000 |
| VŠB - Technical University of Ostrava               | 2150 | 301        | VŠB - Technical University of Ostrava             | Czech Republic           | 18.8     | 22.3          | 15.7     | 85.5      | 32.100000 |
| Wageningen University and Research Center           | 143  | 144        | Wageningen University and Research Center         | Netherlands              | 58.5     | 24.3          | 48.8     | 53.0      | 44.500000 |
| Wake Forest University                              | 91   | 90         | Wake Forest University                            | United States of America | 54.6     | 24.4          | 42.9     | 79.2      | 51.050000 |
| Warsaw University of Technology                     | 2591 | 601        | Warsaw University of Technology                   | Poland                   | 19.4     | 20.7          | 8.5      | 40.3      | 47.400000 |
| Waseda University                                   | 599  | 350        | Waseda University                                 | Japan                    | 25.4     | 27.1          | 17.3     | 29.7      | 27.300000 |
| Washington State University                         | 549  | 301        | Washington State University                       | United States of America | 28.8     | 31.0          | 24.5     | 33.5      | 34.300000 |
| Wayne State University                              | 475  | 251        | Wayne State University                            | United States of America | 34.7     | 24.6          | 16.4     | 55.2      | 32.066667 |
| West University of Timișoara                        | 2595 | 601        | West University of Timișoara                      | Romania                  | 16.1     | 21.0          | 3.9      | 22.4      | 15.766667 |
| Western Sydney University                           | 1800 | 351        | Western Sydney University                         | Australia                | 17.8     | 50.1          | 22.4     | 58.7      | 30.400000 |
| William & Mary                                      | 74   | 75         | William & Mary                                    | United States of America | 53.1     | 20.9          | 36.1     | 95.6      | 53.250000 |
| Wuhan University of Technology                      | 1353 | 301        | Wuhan University of Technology                    | China                    | 14.8     | 18.9          | 7.8      | 78.1      | 58.700000 |
| Xiamen University                                   | 2302 | 401        | Xiamen University                                 | China                    | 26.7     | 25.0          | 15.6     | 47.0      | 29.200000 |
| Xidian University                                   | 2597 | 601        | Xidian University                                 | China                    | 17.9     | 12.8          | 12.1     | 8.9       | 83.700000 |
| Xi'an Jiaotong University                           | 2401 | 501        | Xi'an Jiaotong University                         | China                    | 28.7     | 25.8          | 22.5     | 25.5      | 70.400000 |
| Yale University                                     | 9    | 10         | Yale University                                   | United States of America | 92.1     | 59.2          | 89.7     | 91.5      | 82.475000 |
| Yeshiva University                                  | 69   | 68         | Yeshiva University                                | United States of America | 63.5     | 53.3          | 46.7     | 74.4      | 58.950000 |
| Yeungnam University                                 | 2598 | 601        | Yeungnam University                               | South Korea              | 18.6     | 24.3          | 10.9     | 26.5      | 35.400000 |
| Yokohama National University                        | 2601 | 601        | Yokohama National University                      | Japan                    | 20.1     | 23.3          | 16.0     | 13.5      | 40.400000 |

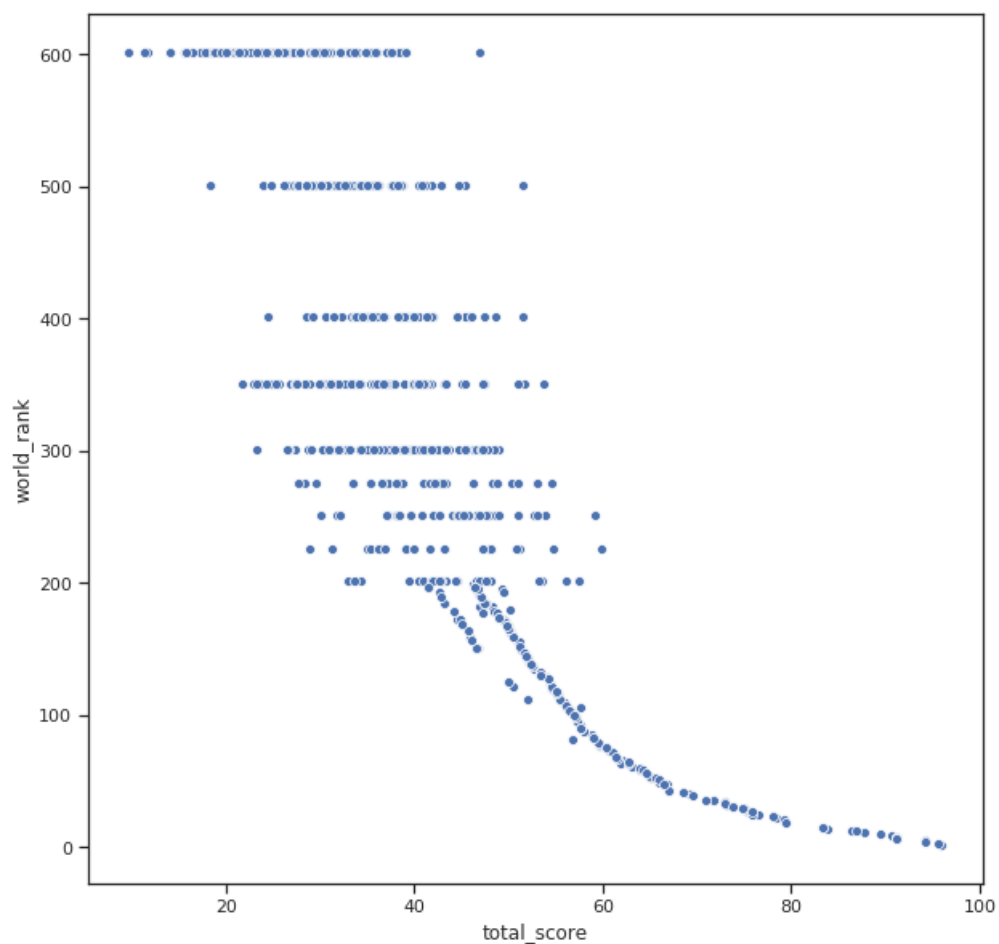
|                                          | world_rank | university_name | country                                  | teaching    | international | research | citations | income |           |
|------------------------------------------|------------|-----------------|------------------------------------------|-------------|---------------|----------|-----------|--------|-----------|
| university_name                          |            |                 |                                          |             |               |          |           |        |           |
| York University                          | 500        | 276             | York University                          | Canada      | 19.9          | 57.7     | 27.7      | 41.2   | 41.700000 |
| Yuan Ze University                       | 601        | 350             | Yuan Ze University                       | Taiwan      | 10.8          | 12.8     | 9.6       | 58.3   | 29.200000 |
| Yildiz Technical University              | 2599       | 601             | Yildiz Technical University              | Turkey      | 14.5          | 14.9     | 7.6       | 19.3   | 44.000000 |
| Zhejiang University                      | 197        | 197             | Zhejiang University                      | China       | 54.6          | 29.6     | 41.3      | 44.3   | 70.300000 |
| École Normale Supérieure                 | 41         | 42              | École Normale Supérieure                 | France      | 66.8          | 44.9     | 48.2      | 95.7   | 30.700000 |
| École Normale Supérieure de Lyon         | 99         | 100             | École Normale Supérieure de Lyon         | France      | 51.1          | 37.6     | 34.4      | 88.8   | 26.100000 |
| École Polytechnique                      | 38         | 39              | École Polytechnique                      | France      | 57.9          | 77.9     | 56.1      | 91.4   | 73.725000 |
| École Polytechnique Fédérale de Lausanne | 47         | 48              | École Polytechnique Fédérale de Lausanne | Switzerland | 55.0          | 100.0    | 56.1      | 83.8   | 38.000000 |
| Örebro University                        | 2134       | 301             | Örebro University                        | Sweden      | 18.3          | 39.4     | 10.4      | 87.8   | 29.800000 |

739 rows × 21 columns



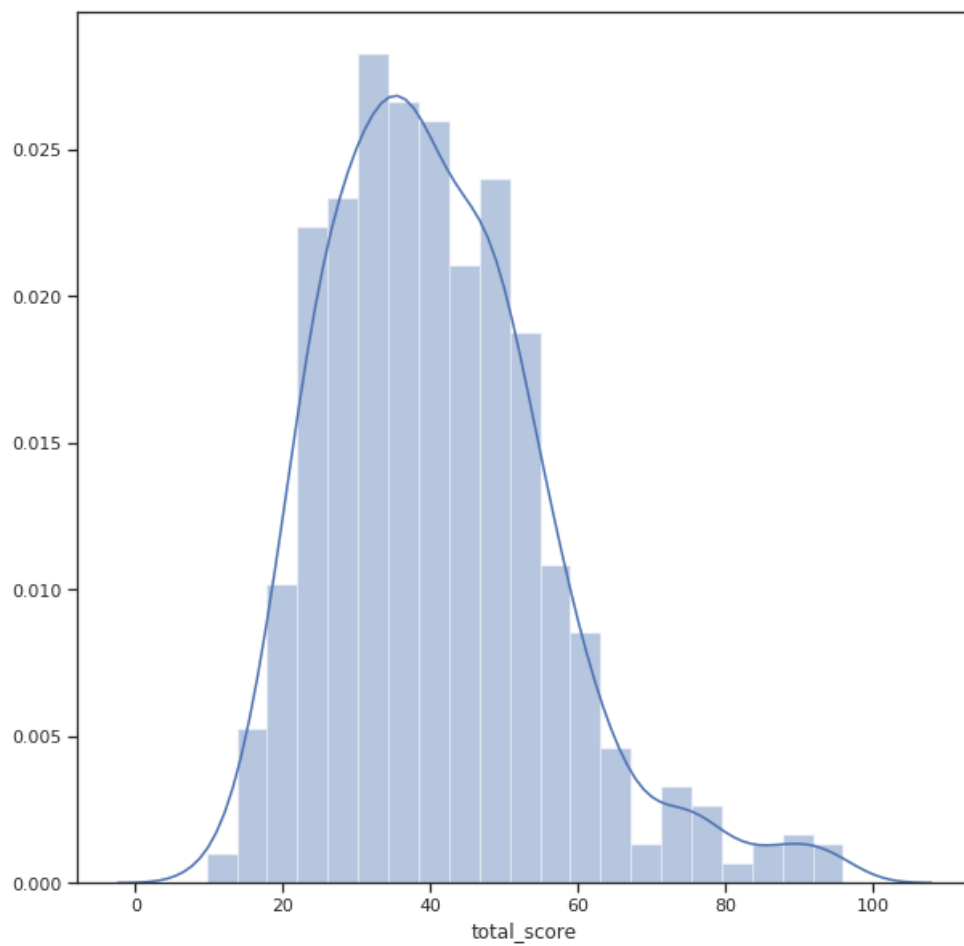
```
In [17]: fig, ax = plt.subplots(figsize=(10,10))
sns.scatterplot(ax=ax, x='total_score', y='world_rank', data=data)
```

```
Out[17]: <matplotlib.axes._subplots.AxesSubplot at 0x7f5cee250ef0>
```



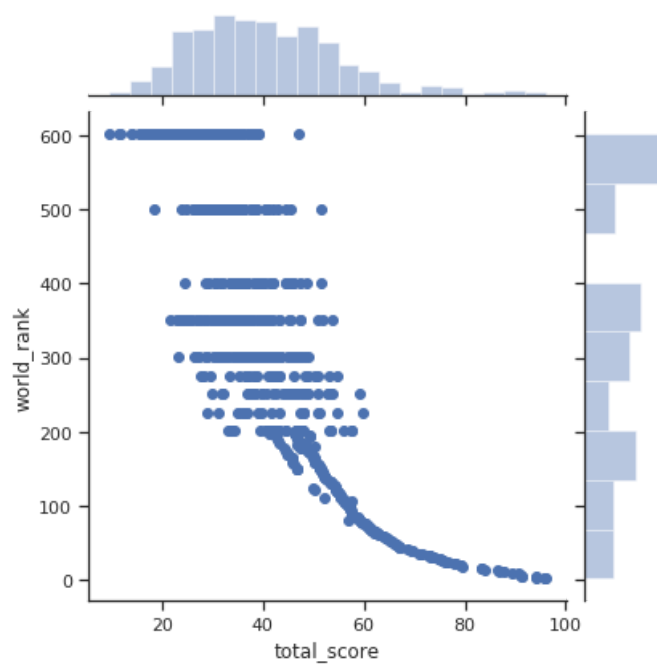
```
In [18]: fig, ax = plt.subplots(figsize=(10,10))
sns.distplot(data['total_score'])
```

```
Out[18]: <matplotlib.axes._subplots.AxesSubplot at 0x7f5cee0257b8>
```



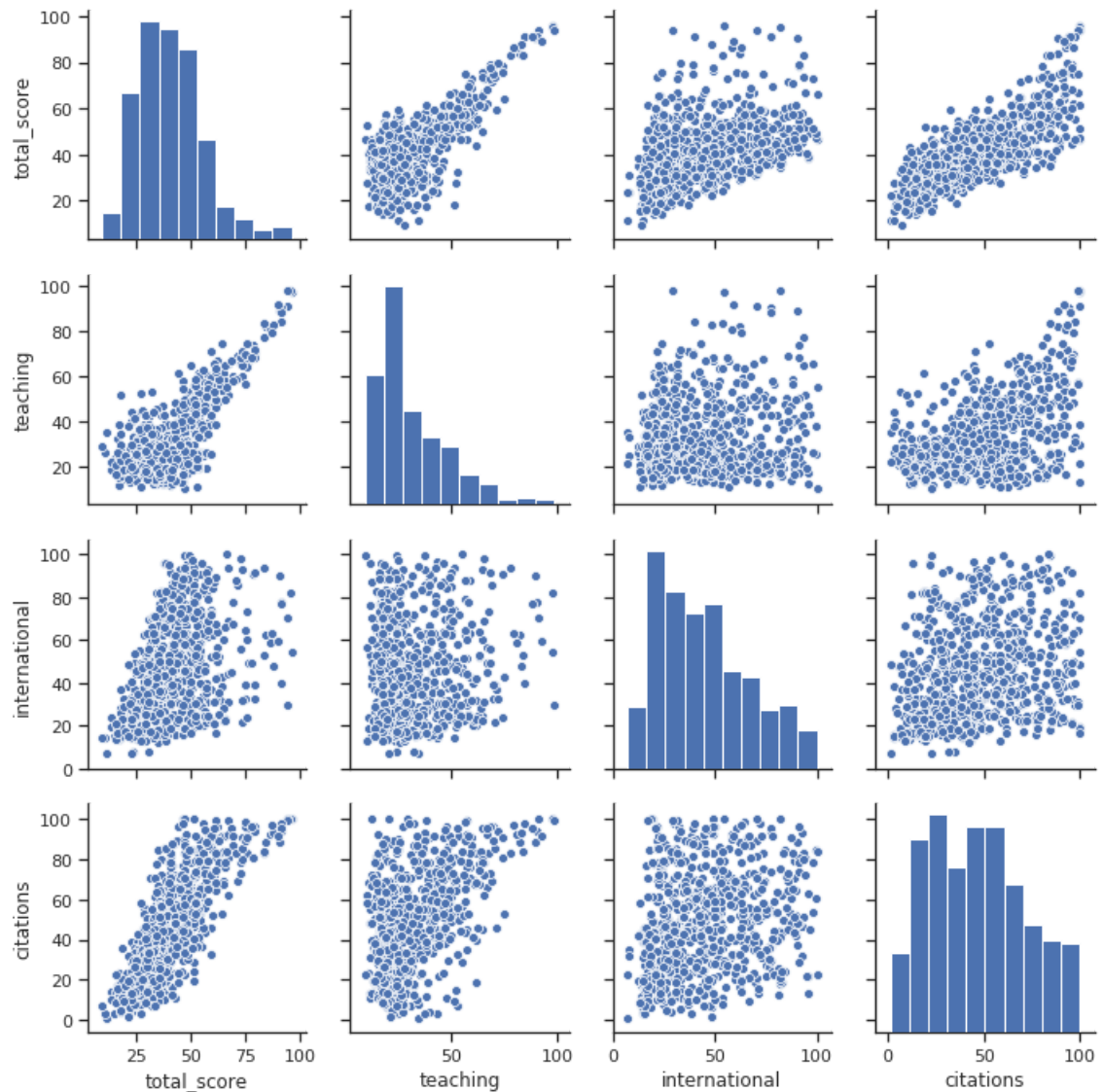
```
In [19]: sns.jointplot(x='total_score', y='world_rank', data=data)
```

```
Out[19]: <seaborn.axisgrid.JointGrid at 0x7f5d21f0e908>
```



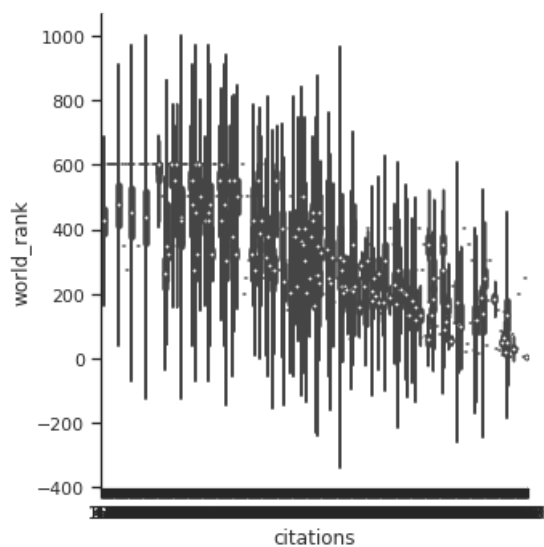
```
In [20]: sns.pairplot(data[['total_score', 'teaching', 'international', 'citations']])
```

```
Out[20]: <seaborn.axisgrid.PairGrid at 0x7f5ceddd1e80>
```



```
In [21]: sns.catplot(y='world_rank', x='citations', data=data, kind="violin", split=True)
```

```
Out[21]: <seaborn.axisgrid.FacetGrid at 0x7f5cedf75fd0>
```



## 4) Информация о корреляции признаков

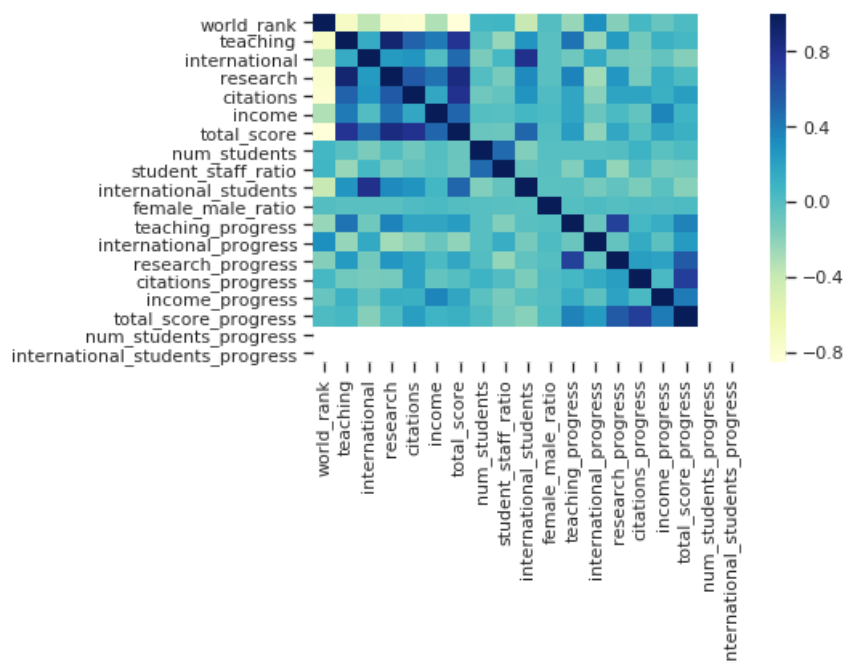
```
In [22]: data.corr()
```

```
Out[22]:
```

|                                 | world_rank | teaching  | international | research  | citations | income    | total_score | num_st |
|---------------------------------|------------|-----------|---------------|-----------|-----------|-----------|-------------|--------|
| world_rank                      | 1.000000   | -0.732092 | -0.365120     | -0.785236 | -0.805942 | -0.317702 | -0.853945   | 0.0    |
| teaching                        | -0.732092  | 1.000000  | 0.143880      | 0.908191  | 0.521600  | 0.408316  | 0.770419    | -0.0   |
| international                   | -0.365120  | 0.143880  | 1.000000      | 0.240472  | 0.273609  | 0.013415  | 0.484897    | -0.0   |
| research                        | -0.785236  | 0.908191  | 0.240472      | 1.000000  | 0.561164  | 0.444592  | 0.846262    | 0.0    |
| citations                       | -0.805942  | 0.521600  | 0.273609      | 0.561164  | 1.000000  | 0.167780  | 0.792220    | -0.0   |
| income                          | -0.317702  | 0.408316  | 0.013415      | 0.444592  | 0.167780  | 1.000000  | 0.498773    | -0.0   |
| total_score                     | -0.853945  | 0.770419  | 0.484897      | 0.846262  | 0.792220  | 0.498773  | 1.000000    | -0.0   |
| num_students                    | 0.065192   | -0.020426 | -0.137339     | 0.005305  | -0.099872 | -0.005632 | -0.075964   | 1.0    |
| student_staff_ratio             | 0.080110   | -0.232529 | 0.056522      | -0.136723 | -0.055355 | -0.008549 | -0.087236   | 0.0    |
| international_students          | -0.393047  | 0.275345  | 0.803313      | 0.339426  | 0.279234  | 0.058790  | 0.503380    | -0.0   |
| female_male_ratio               | 0.003110   | -0.018707 | -0.014194     | -0.018145 | 0.016426  | 0.033576  | 0.005745    | -0.0   |
| teaching_progress               | -0.231821  | 0.438598  | -0.096220     | 0.365263  | 0.177647  | 0.177694  | 0.236082    | -0.0   |
| international_progress          | 0.308062   | -0.225894 | 0.156078      | -0.268961 | -0.189165 | -0.072636 | -0.205629   | -0.0   |
| research_progress               | -0.177708  | 0.246676  | -0.109139     | 0.267249  | 0.192014  | 0.035306  | 0.172600    | 0.0    |
| citations_progress              | 0.056633   | -0.102778 | -0.131552     | -0.118460 | 0.191262  | -0.054486 | -0.004056   | 0.0    |
| income_progress                 | -0.065488  | 0.108176  | -0.059840     | 0.123628  | 0.110871  | 0.361134  | 0.177728    | -0.0   |
| total_score_progress            | 0.023829   | 0.056909  | -0.179233     | 0.027362  | 0.219279  | 0.082414  | 0.114138    | 0.0    |
| num_students_progress           | NaN        | NaN       | NaN           | NaN       | NaN       | NaN       | NaN         |        |
| international_students_progress | NaN        | NaN       | NaN           | NaN       | NaN       | NaN       | NaN         |        |

```
In [23]: sns.heatmap(data.corr(), cmap='YlGnBu', fmt='.0f')
```

```
Out[23]: <matplotlib.axes._subplots.AxesSubplot at 0x7f5ce530b278>
```



```
In [ ]:
```