# Description

The 'Olympics' is a global athletic event with roots that trace back to ancient Greece and continue throughout the better part of the Roman period. Modern Olympics were revived in the 19th century and have been held every 4 years since 1896 in different cities around the world. There is competition among nations to host the Olympic Games and competition among athletes to win in their respective sports. Both of these feats are considered to be very prestigious. In this context, it would be interesting to study a dataset of 'olympic winners' and discover the footprint of humanity's modern history on it. On a high level, we will be collecting and working with an Olympic Games dataset that contains medals won per country. We will be reading and parsing these values from Wikipedia, performing small computations on the accumulated data and finally formatting and outputting some values in a .csv file.

# Overall Design

The solution should be organised as a typical RPA process, using a single Work Queue (Blue prism and UiPath) or a data table (Automation Anywhere). Each Work Queue or data table item should correspond to each year since 1896 (included) until today

# Wikipedia

For our purposes we will be 'scraping' our dataset from Wikipedia. We are only interested in the Summer Olympics. For every year that the modern era Games were scheduled (e.g. 1896, 1940, 1984) Wikipedia has a page about it with similar format. If the Games were not disrupted, the page contains a table with the top performing nations for that year, as displayed below.

## Medal count [ edit ]

*Main article: 1992 Summer Olympics medal table*

The following table reflects the top ten nations in terms of total medals won at the 1992 Games (the host nation is highlighted).

| Rank | Nation | Gold | Silver | Bronze | Total |
|---|---|---|---|---|---|
| 1 | Unified Team | 45 | 38 | 29 | 112 |
| 2 | United States | 37 | 34 | 37 | 108 |
| 3 | Germany | 33 | 21 | 28 | 82 |
| 4 | China | 16 | 22 | 16 | 54 |
| 5 | Cuba | 14 | 6 | 11 | 31 |
| 6 | Spain* | 13 | 7 | 2 | 22 |
| 7 | South Korea | 12 | 5 | 12 | 29 |
| 8 | Hungary | 11 | 12 | 7 | 30 |
| 9 | France | 8 | 5 | 16 | 29 |
| 10 | Australia | 7 | 9 | 11 | 27 |
| **Totals (10 nations)** | | **196** | **159** | **169** | **524** |

We would like to 'scrape' this table and read it into Blue Prism, UiPath or Automation Anywhere. Inside your data structure you should retain the fields: 'Nation', 'Gold', 'Silver', 'Bronze' for every row in the Wikipedia table.

# REST Countries API

Subsequently, we will match every country to a 'subregion'. To do that we will make use of the publicly available HTTP API 'REST Countries'. API Reference
Provide the 'name' of the respective countries, get and parse the 'subregion' from the response and store it inside Blue Prism as part of your solution.

## Output File

Create a .csv file with the nation/country that won the most silver medals each year:

- The first row should be a header with the columns ‘nation’, ‘subregion’, ‘#silver’
- Create a row in the .csv for every Work Queue item with the data indicated by the header