

Labour and Employment

By Gerard Utoware

Table of Contents

Overviews – Page 3

Q1 – Page 4

Overview of the tool used

The main advantage of python in comparison to other data analytic tools such as excel, would be Python's ability to handle large volumes of data without hindering productivity. Running scripts using the libraries on allows for more automation and ease of obtaining analysed data. With this task, the data was imported from python which ensures that data will not be lost or tampered with while performing data analysis. After defining the columns from the imported data, it was less tedious to run scripts than to perform individual analysis on excel.

Overview of the libraries used

Numpy, pandas and matplotlib were used in the script. All three python libraries allow different functions to be performed. Pandas allows data analysis scripts to be conducted on python. It utilises two other libraries, being numpy and matplotlib. Matplotlib allows for data visualisation through a variety of charts such as bar and pie charts. Numpy permits mathematical operations in python. The combination of these three libraries is what allowed the following data and charts to be formed and analysed. In addition to these, seaborn was installed in order for a heatmap to be printed and shown.

Overview of the dataset

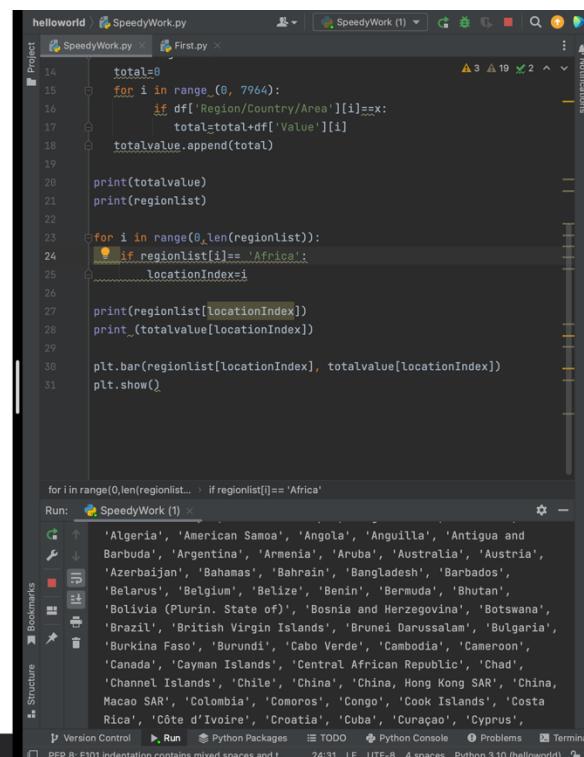
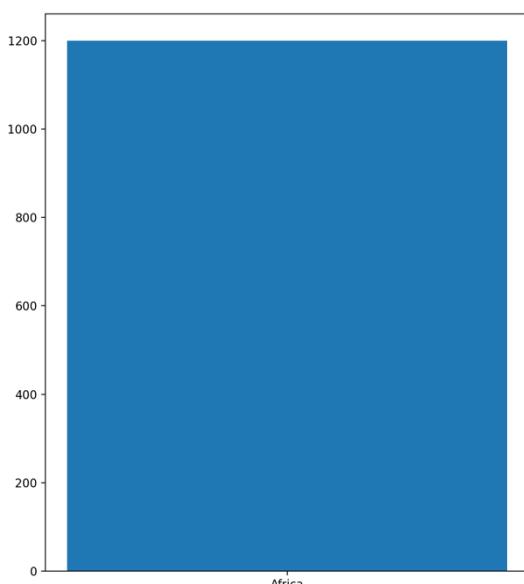
The dataset imported onto python sought to tabulate the employment rates of people in different countries/regions/areas across different years and different industries provided by sources such as the International Labour Organization (ILO) and Geneva. The python script looked at the sum of these rates by 'countries/regions/areas' in order to highlight which regions had the lowest rates of total employment. What the data eventually told was that majority of the countries/regions/areas had a sum value of approximately 1200. This tells us that this is the region of good employment rates. With 36 total entries per countries/regions/areas, the average of a 'good' employment rate per year and industry would be 33.3%. Countries/regions/areas that fall below this mark will be deemed as outliers and will be below par in this regard. The violin plot, scatter graph and boxplot in particular will aid in displaying the outliers, however, due to the sheer number of countries/regions/areas, naming these outliers will be difficult.

Q1

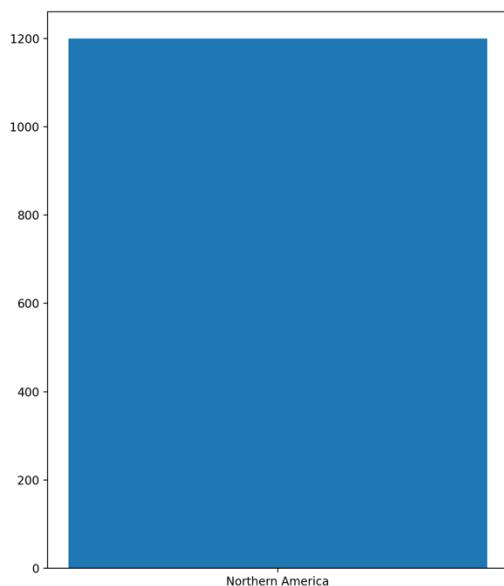
The screenshot shows a Python development environment with the following details:

- Project:** A project named "SpeedyWork" is open.
- Code Editor:** The file "SpeedyWork.py" contains Python code to calculate total values for regions and create a bar chart for Nigeria. The code includes imports for pandas and matplotlib, defines a function to calculate totals, and uses a bar chart to visualize the results.
- Run Tab:** The "Run" tab shows the output of the script, which lists various countries and their total values, followed by the specific output for "Nigeria": "Nigeria" and "1199.999999999998".
- Output Tab:** The "Python Console" tab shows the command "plt.show()" being run.
- Result:** A bar chart titled "SpeedyWork ()" is displayed on the right, showing a single blue bar for "Nigeria" reaching a value of approximately 1199.999999999998.

Bar chart depicting the sum of the 'Value' column per region/country/area, across all years. In this case being Nigeria.



Bar chart depicting the sum of the 'Value' column per region/country/area, across all years. In this case being Africa.

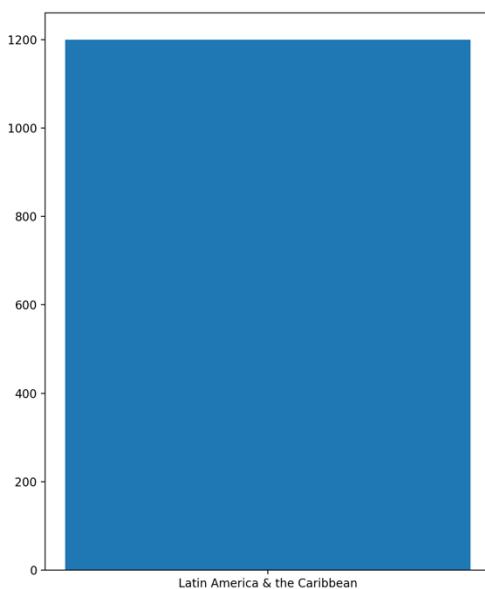


```

helloworld SpeedyWork.py First.py
Project SpeedyWork.py
14     total=0
15     for i in range_(0, 7964):
16         if df['Region/Country/Area'][i]==x:
17             total=total+df['Value'][i]
18
19     print(totalvalue)
20     print(regionlist)
21
22     for i in range(0,len(regionlist)):
23         if regionlist[i]== 'Northern America':
24             locationIndex=i
25         #elif regionlist[i] == 'Americas':
26         #    locationIndex2 = i
27
28         print(regionlist[locationIndex])
29         print_(totalvalue[locationIndex])
30         #print(regionlist[locationIndex2])
31         #print_(totalvalue[locationIndex2])
32
33     plt.bar(regionlist[locationIndex], totalvalue[locationIndex])
34     plt.show()
35
for i in range(0,len(regionlist)): if regionlist[i]== 'Northern Am...
Run: SpeedyWork () ×
Palestine', 'Sudan', 'Suriname', 'Sweden', 'Switzerland', 'Syrian Arab Republic', 'Tajikistan', 'Thailand', 'Timor-Leste', 'Togo', 'Tonga', 'Trinidad and Tobago', 'Tunisia', 'Turkey', 'Turkmenistan', 'Turks and Caicos Islands', 'Uganda', 'Ukraine', 'United Arab Emirates', 'United Kingdom', 'United Rep. of Tanzania', 'United States of America', 'United States Virgin Islands', 'Uruguay', 'Uzbekistan', 'Vanuatu', 'Venezuela (Boliv. Rep. of)', 'Viet Nam', 'Western Sahara', 'Yemen', 'Zambia', 'Zimbabwe', 'European Union (EU)']
Northern America
1199.9999999999998

```

Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Northern America.

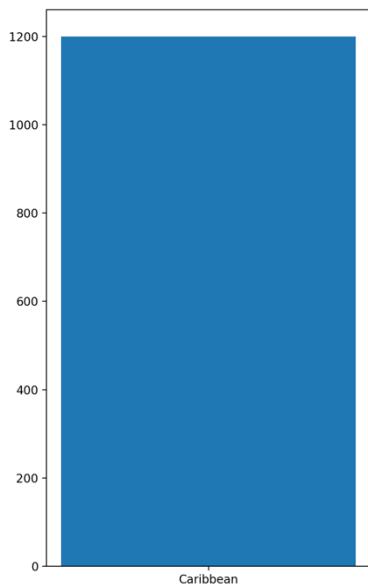


```

helloworld SpeedyWork.py First.py
Project SpeedyWork.py
14     total=0
15     for i in range_(0, 7964):
16         if df['Region/Country/Area'][i]==x:
17             total=total+df['Value'][i]
18
19     print(totalvalue)
20     print(regionlist)
21
22     for i in range(0,len(regionlist)):
23         if regionlist[i]== 'Latin America & the Caribbean':
24             locationIndex=i
25
26         print(regionlist[locationIndex])
27         print_(totalvalue[locationIndex])
28
29
30     plt.bar(regionlist[locationIndex], totalvalue[locationIndex])
31     plt.show()
32
for i in range(0,len(regionlist)): if regionlist[i]== 'Latin Ameri...
Run: SpeedyWork () ×
Spain', 'Sri Lanka', 'State of Palestine', 'Sudan', 'Suriname', 'Sweden', 'Switzerland', 'Syrian Arab Republic', 'Tajikistan', 'Thailand', 'Timor-Leste', 'Togo', 'Tonga', 'Trinidad and Tobago', 'Tunisia', 'Turkey', 'Turkmenistan', 'Turks and Caicos Islands', 'Uganda', 'Ukraine', 'United Arab Emirates', 'United Kingdom', 'United Rep. of Tanzania', 'United States of America', 'United States Virgin Islands', 'Uruguay', 'Uzbekistan', 'Vanuatu', 'Venezuela (Boliv. Rep. of)', 'Viet Nam', 'Western Sahara', 'Yemen', 'Zambia', 'Zimbabwe', 'European Union (EU)']
Latin America & the Caribbean
1200.0000000000001

```

Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Latin America & the Caribbean.



```
total=0
for i in range(0, 7964):
    if df['Region/Country/Area'][i]==x:
        total+=df['Value'][i]
totalvalue.append(total)

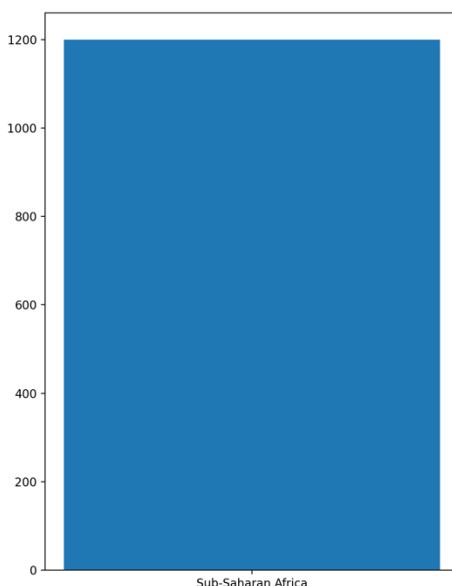
print(totalvalue)
print(regionlist)

for i in range(0,len(regionlist)):
    if regionlist[i]=='Caribbean':
        locationIndex=i

print(regionlist[locationIndex])
print_(totalvalue[locationIndex])

plt.bar(regionlist[locationIndex], totalvalue[locationIndex])
plt.show()
```

Bar chart depicting the sum of the 'Value' column per region/country/area, across all years. In this case being the Caribbean.

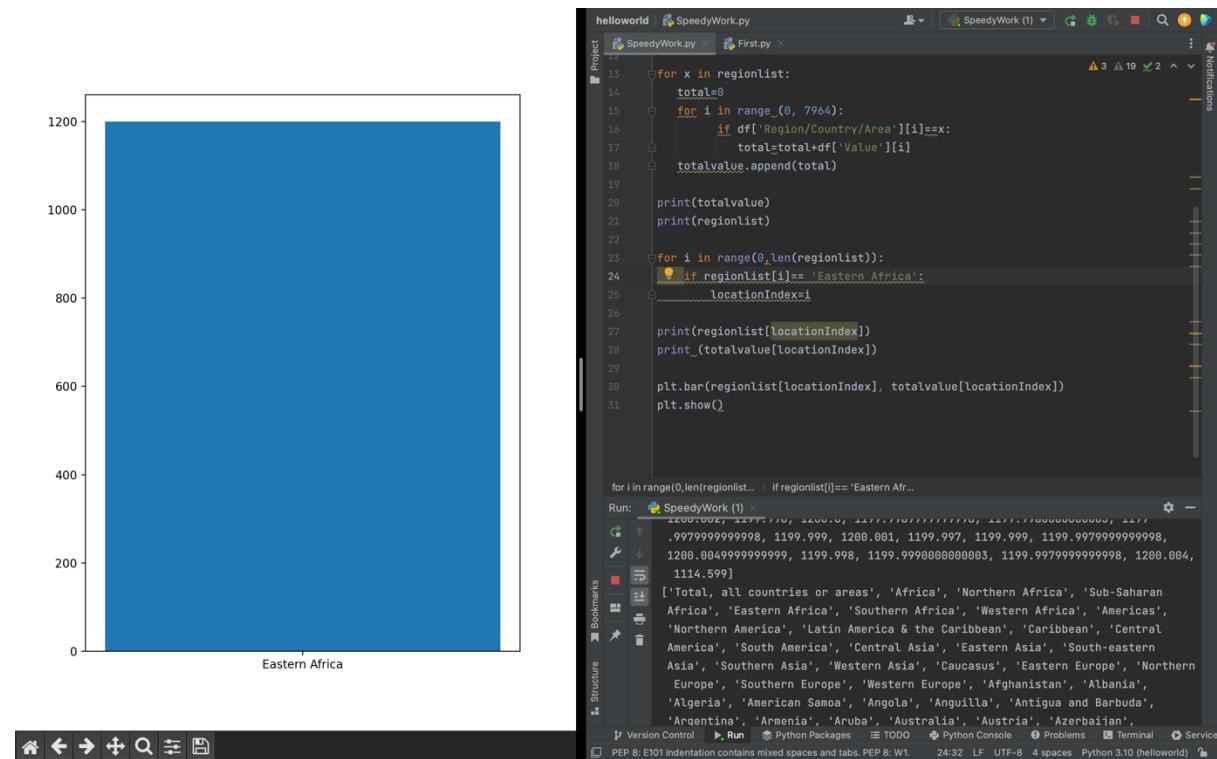


The screenshot shows the PyCharm IDE interface with the following details:

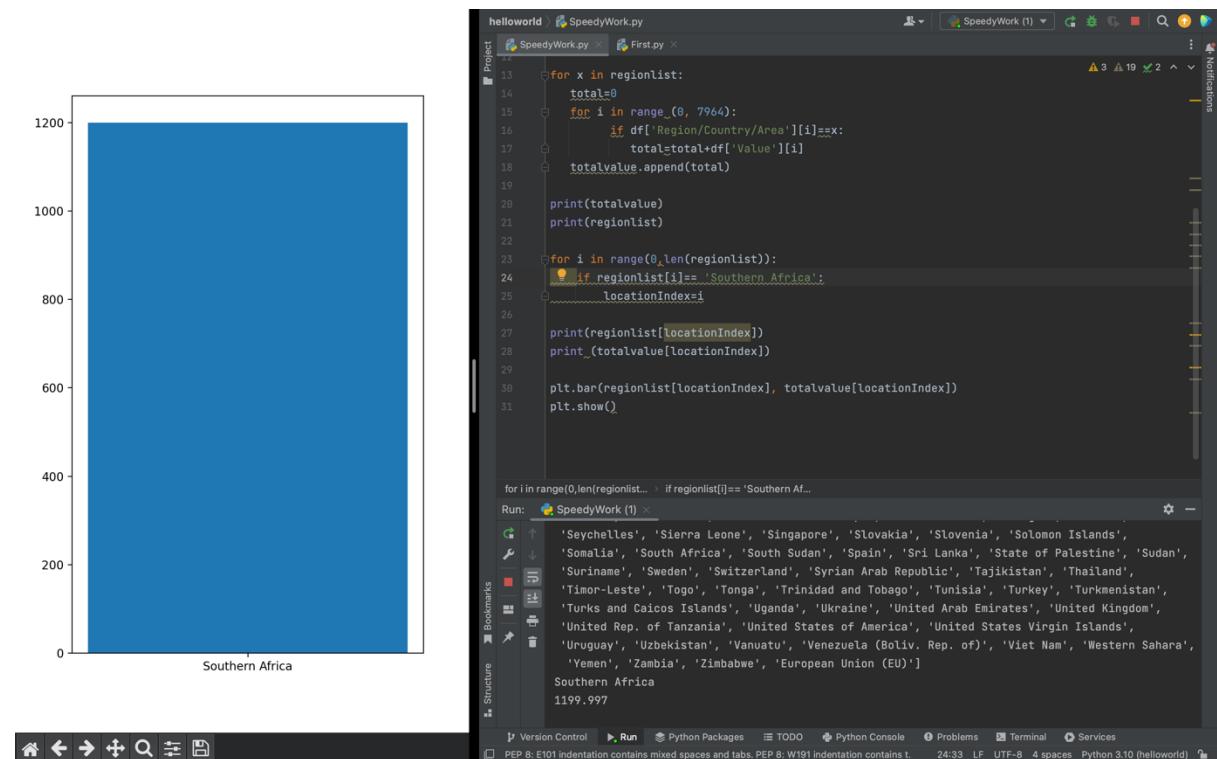
- Project Bar:** Shows 'helloworld' and 'SpeedyWork.py'.
- Code Editor:** Displays the 'SpeedyWork.py' file content. The code uses pandas to read a CSV file, filter regions, and calculate total values. A specific line for 'Sub-Saharan Africa' is highlighted with a yellow background and a lightbulb icon, indicating a potential issue or suggestion.
- Run Bar:** Shows 'SpeedyWork (1)'.
- Sidebar:** Includes 'Bookmarks' and 'Structure' tabs.
- Bottom Navigation:** Includes 'Version Control', 'Run', 'Python Packages', 'TODO', 'Python Console', 'Problems', 'Terminal', and 'Services'.

```
for x in regionlist:  
    total=0  
    for i in range(0, 7964):  
        if df['Region/Country/Area'][i]==x:  
            total+=df['Value'][i]  
    totalvalue.append(total)  
  
print(totalvalue)  
print(regionlist)  
  
for i in range(0,len(regionList)):  
    if regionList[i]=='Sub-Saharan Africa':  
        locationIndex=i  
  
    print(regionList[locationIndex])  
    print_(totalvalue[locationIndex])  
  
plt.bar(regionlist[locationIndex], totalvalue[locationIndex])  
plt.show()
```

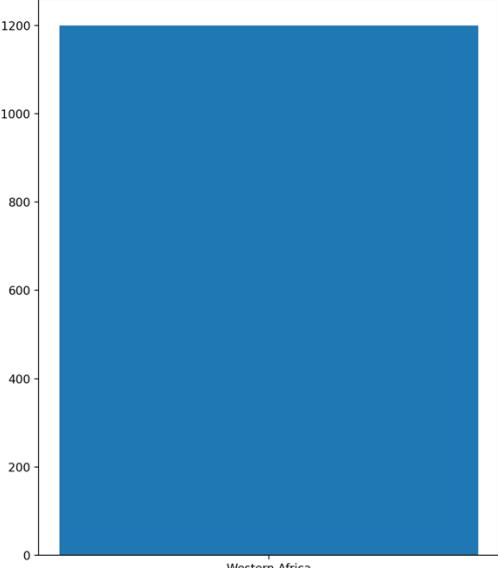
Bar chart depicting the sum of the 'Value' column per region/country/area, across all years. In this case being Sub-Saharan Africa.



Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Eastern Africa.



Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Southern Africa.



Western Africa

```

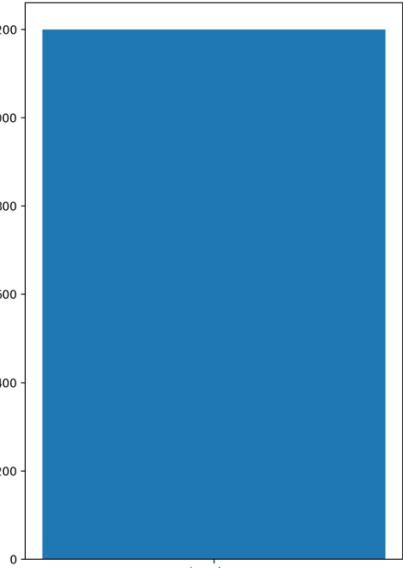
helloworld > SpeedyWork.py First.py
SpeedyWork.py First.py
13     for x in regionlist:
14         total=0
15         for i in range(0, 7964):
16             if df['Region/Country/Area'][i]==x:
17                 total=total+df['Value'][i]
18         totalvalue.append(total)
19
20     print(totalvalue)
21     print(regionlist)
22
23     for i in range(0,len(regionlist)):
24         if regionlist[i]=='Western Africa':
25             locationIndex=i
26
27     print(regionlist[locationIndex])
28     print_(totalvalue[locationIndex])
29
30     plt.bar(regionlist[locationIndex], totalvalue[locationIndex])
31     plt.show()

for i in range(0,len(regionlist)): if regionlist[i]=='Western Afr...
Run: SpeedyWork (1) ×
Palestine', 'Sudan', 'Suriname', 'Sweden', 'Switzerland', 'Syrian Arab Republic', 'Tajikistan', 'Thailand', 'Timor-Leste', 'Togo', 'Tonga', 'Trinidad and Tobago', 'Tunisia', 'Turkey', 'Turkmenistan', 'Turks and Caicos Islands', 'Uganda', 'Ukraine', 'United Arab Emirates', 'United Kingdom', 'United Rep. of Tanzania', 'United States of America', 'United States Virgin Islands', 'Uruguay', 'Uzbekistan', 'Vanuatu', 'Venezuela (Boliv. Rep. of)', 'Viet Nam', 'Western Sahara', 'Yemen', 'Zambia', 'Zimbabwe', 'European Union (EU)']
Western Africa
1200.002

```

Version Control Run Python Packages TODO Python Console Problems Terminal Services

PEP 8: E101 indentation contains mixed spaces and tabs, PEP 8: W191 indentation co... 24-32 LF UTF-8 4 spaces Python 3.10 (helloworld)



Americas

```

helloworld > SpeedyWork.py First.py
SpeedyWork.py First.py
13     for x in regionlist:
14         total=0
15         for i in range(0, 7964):
16             if df['Region/Country/Area'][i]==x:
17                 total=total+df['Value'][i]
18         totalvalue.append(total)
19
20     print(totalvalue)
21     print(regionlist)
22
23     for i in range(0,len(regionlist)):
24         if regionlist[i]=='Americas':
25             locationIndex=i
26
27     print(regionlist[locationIndex])
28     print_(totalvalue[locationIndex])
29
30     plt.bar(regionlist[locationIndex], totalvalue[locationIndex])
31     plt.show()

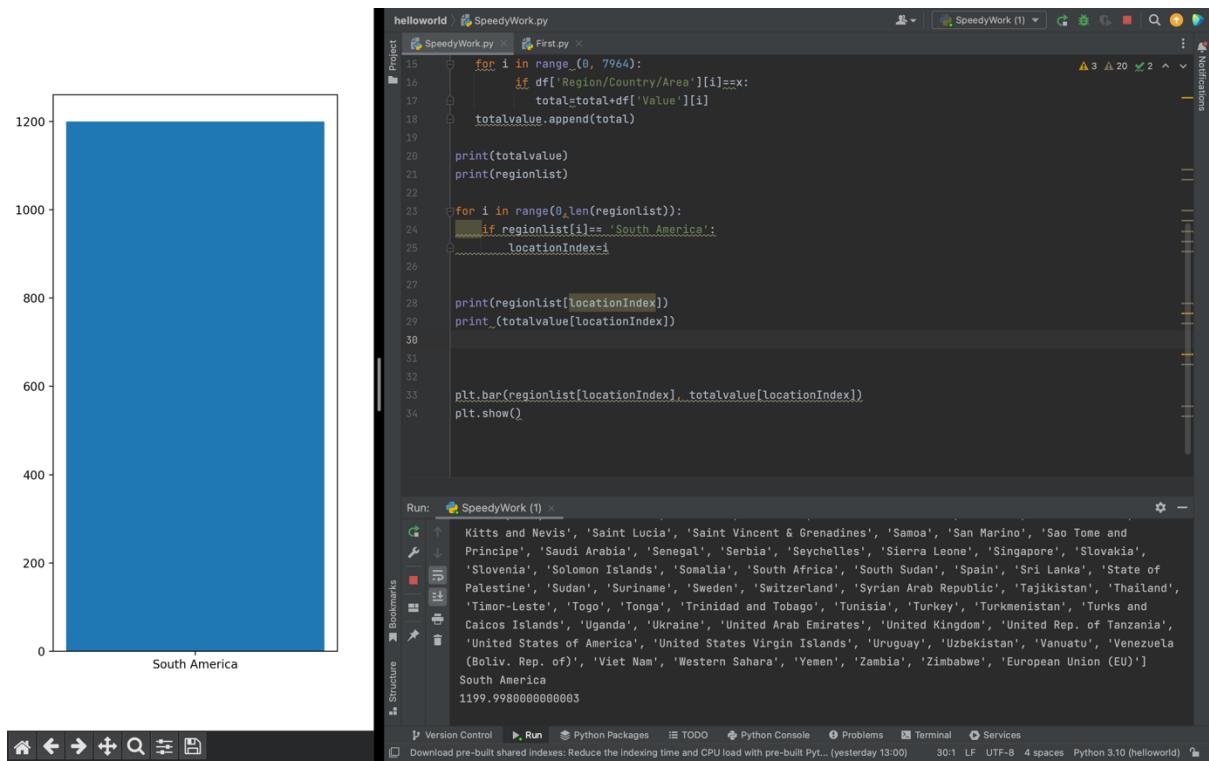
for i in range(0,len(regionlist)): if regionlist[i]=='Americas':
Run: SpeedyWork (1) ×
'Sierra Leone', 'Singapore', 'Slovakia', 'Slovenia', 'Solomon Islands', 'Somalia', 'South Africa', 'South Sudan', 'Spain', 'Sri Lanka', 'State of Palestine', 'Sudan', 'Suriname', 'Sweden', 'Switzerland', 'Syrian Arab Republic', 'Tajikistan', 'Thailand', 'Timor-Leste', 'Togo', 'Tonga', 'Trinidad and Tobago', 'Tunisia', 'Turkey', 'Turkmenistan', 'Turks and Caicos Islands', 'Uganda', 'Ukraine', 'United Arab Emirates', 'United Kingdom', 'United Rep. of Tanzania', 'United States of America', 'United States Virgin Islands', 'Uruguay', 'Uzbekistan', 'Vanuatu', 'Venezuela (Boliv. Rep. of)', 'Viet Nam', 'Western Sahara', 'Yemen', 'Zambia', 'Zimbabwe', 'European Union (EU)']
Americas
1200.0

```

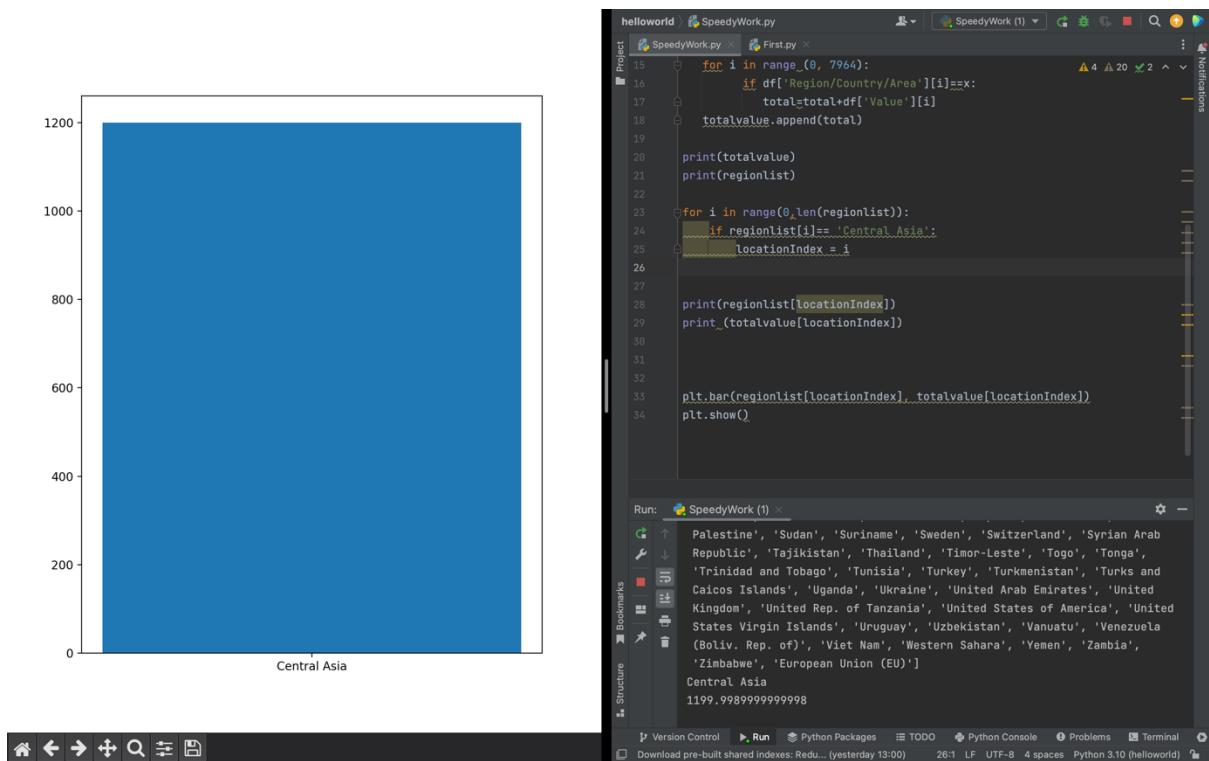
Version Control Run Python Packages TODO Python Console Problems Terminal Services

PEP 8: E101 indentation contains mixed spaces and tabs, PEP 8: W191 indentation co... 24-33 LF UTF-8 4 spaces Python 3.10 (helloworld)

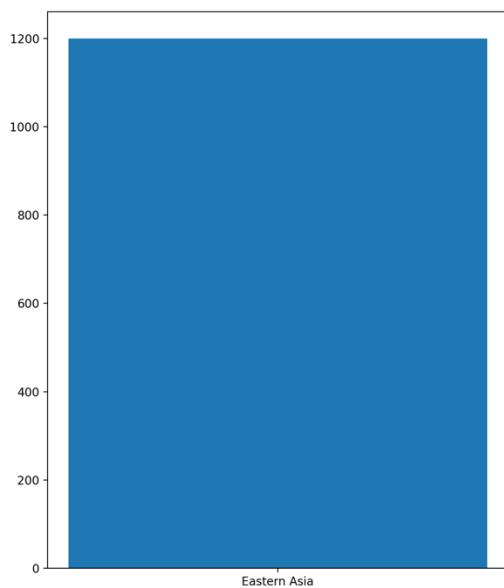
Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being the Americas



Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being South America.



Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Central Asia.



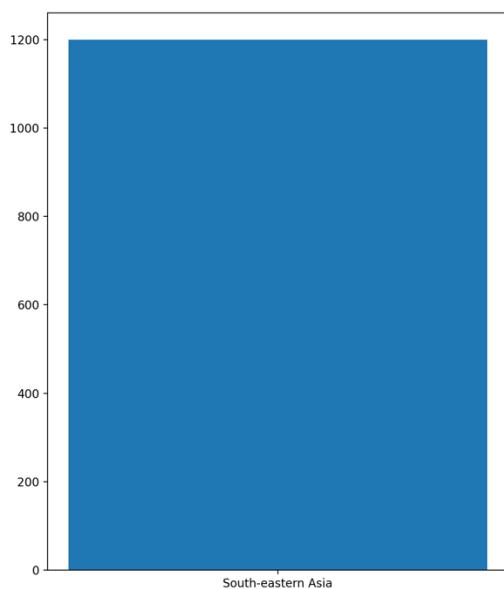
```

helloworld | SpeedyWork.py | Firstpy | SpeedyWork (1) | G | B | S | M | D | N | Notifications
15     for i in range(0, 7964):
16         if df['Region/Country/Area'][i]==x:
17             total=total+df['Value'][i]
18
19     print(totalvalue)
20     print(regionlist)
21
22     for i in range(0,len(regionlist)):
23         if regionlist[i]== 'Eastern_Aisa':
24             locationIndex = i
25
26
27
28     print(regionlist[locationIndex])
29     print_(totalvalue[locationIndex])
30
31
32
33     plt.bar(regionlist[locationIndex],totalvalue[locationIndex])
34     plt.show()

```

Run: SpeedyWork (1) | Version Control | Run | Python Packages | TODO | Python Console | Problems | Terminal | Download pre-built shared indexes: Red... (yesterday 13:00) | 28:33 | LF | UTF-8 | 4 spaces | Python 3.10 (helloworld)

Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Eastern Asia.



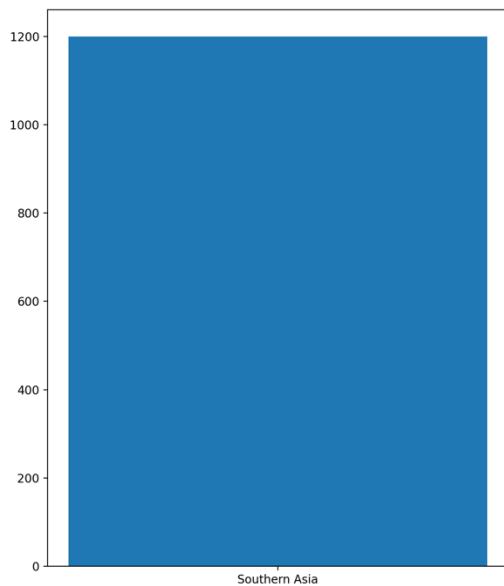
```

helloworld | SpeedyWork.py | Firstpy | SpeedyWork (1) | G | B | S | M | D | N | Notifications
15     for i in range(0, 7964):
16         if df['Region/Country/Area'][i]==x:
17             total=total+df['Value'][i]
18
19     print(totalvalue)
20     print(regionlist)
21
22     for i in range(0,len(regionlist)):
23         if regionlist[i]== 'South-easte...':
24             locationIndex = i
25
26
27
28     print(regionlist[locationIndex])
29     print_(totalvalue[locationIndex])
30
31
32
33     plt.bar(regionlist[locationIndex],totalvalue[locationIndex])
34     plt.show()

```

Run: SpeedyWork (1) | Version Control | Run | Python Packages | TODO | Python Console | Problems | Terminal | Download pre-built shared indexes: Red... (yesterday 13:00) | 24:32 | LF | UTF-8 | 4 spaces | Python 3.10 (helloworld)

Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being South-Eastern Asia.



```

helloworld SpeedyWork.py First.py
SpeedyWork.py
for i in range(0, 7964):
    if df['Region/Country/Area'][i]==x:
        total=total+df['Value'][i]
    totalvalue.append(total)

print(totalvalue)
print(regionlist)

for i in range(0,len(regionlist)):
    if regionlist[i]=='Southern Asia':
        locationIndex = i

print(regionlist[locationIndex])
print_(totalvalue[locationIndex])

plt.bar(regionlist[locationIndex], totalvalue[locationIndex])
plt.show()

```

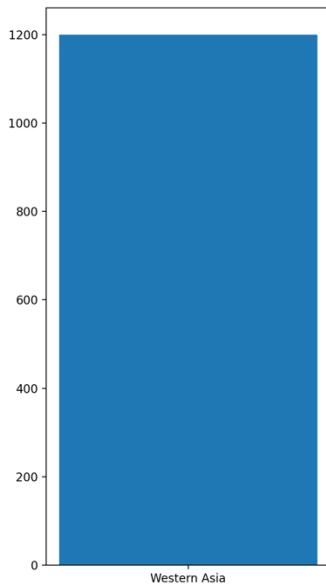
Run: SpeedyWork ()

```

Palestine', 'Sudan', 'Suriname', 'Sweden', 'Switzerland', 'Syrian Arab Republic', 'Tajikistan', 'Thailand', 'Timor-Leste', 'Togo', 'Tonga', 'Trinidad and Tobago', 'Tunisia', 'Turkey', 'Turkmenistan', 'Turks and Caicos Islands', 'Uganda', 'Ukraine', 'United Arab Emirates', 'United Kingdom', 'United Rep. of Tanzania', 'United States of America', 'United States Virgin Islands', 'Uruguay', 'Uzbekistan', 'Vanuatu', 'Venezuela (Boliv. Rep. of)', 'Viet Nam', 'Western Sahara', 'Yemen', 'Zambia', 'Zimbabwe', 'European Union (EU)']
Southern Asia
1199.9989999999998

```

Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Southern Asia.



```

helloworld SpeedyWork.py First.py
SpeedyWork.py
for i in range(0, 7964):
    if df['Region/Country/Area'][i]==x:
        total=total+df['Value'][i]
    totalvalue.append(total)

print(totalvalue)
print(regionlist)

for i in range(0,len(regionlist)):
    if regionlist[i]=='Western Asia':
        locationIndex = i

print(regionlist[locationIndex])
print_(totalvalue[locationIndex])

plt.bar(regionlist[locationIndex], totalvalue[locationIndex])
plt.show()

```

Run: SpeedyWork ()

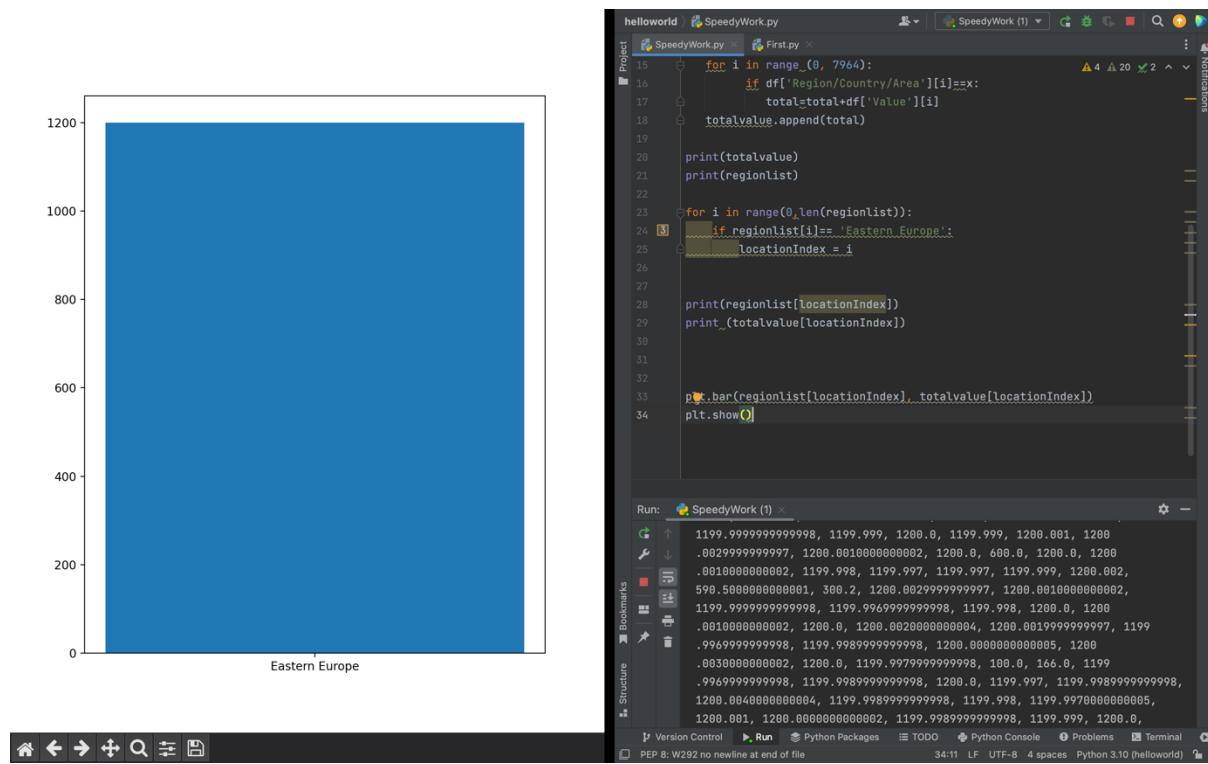
```

Kitts and Nevis', 'Saint Lucia', 'Saint Vincent & Grenadines', 'Samoa', 'San Marino', 'Sao Tome and Principe', 'Saudi Arabia', 'Senegal', 'Serbia', 'Seychelles', 'Sierra Leone', 'Singapore', 'Slovakia', 'Slovenia', 'Solomon Islands', 'Somalia', 'South Africa', 'South Sudan', 'Spain', 'Sri Lanka', 'State of Palestine', 'Sudan', 'Suriname', 'Sweden', 'Switzerland', 'Syrian Arab Republic', 'Tajikistan', 'Thailand', 'Timor-Leste', 'Togo', 'Tonga', 'Trinidad and Tobago', 'Tunisia', 'Turkey', 'Turkmenistan', 'Turks and Caicos Islands', 'Uganda', 'Ukraine', 'United Arab Emirates', 'United Kingdom', 'United Rep. of Tanzania', 'United States of America', 'United States Virgin Islands', 'Uruguay', 'Uzbekistan', 'Vanuatu', 'Venezuela (Boliv. Rep. of)', 'Viet Nam', 'Western Sahara', 'Yemen', 'Zambia', 'Zimbabwe', 'European Union (EU)']
Western Asia
1199.9980000000003

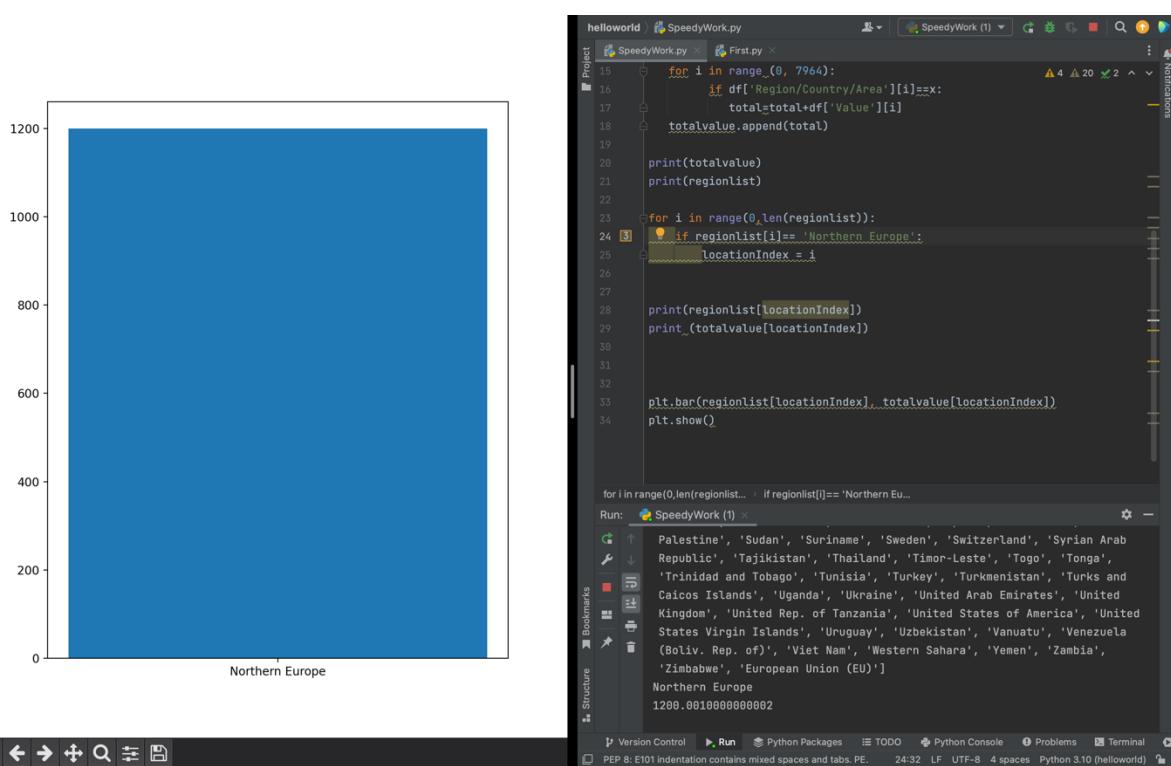
```

Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Western Asia.

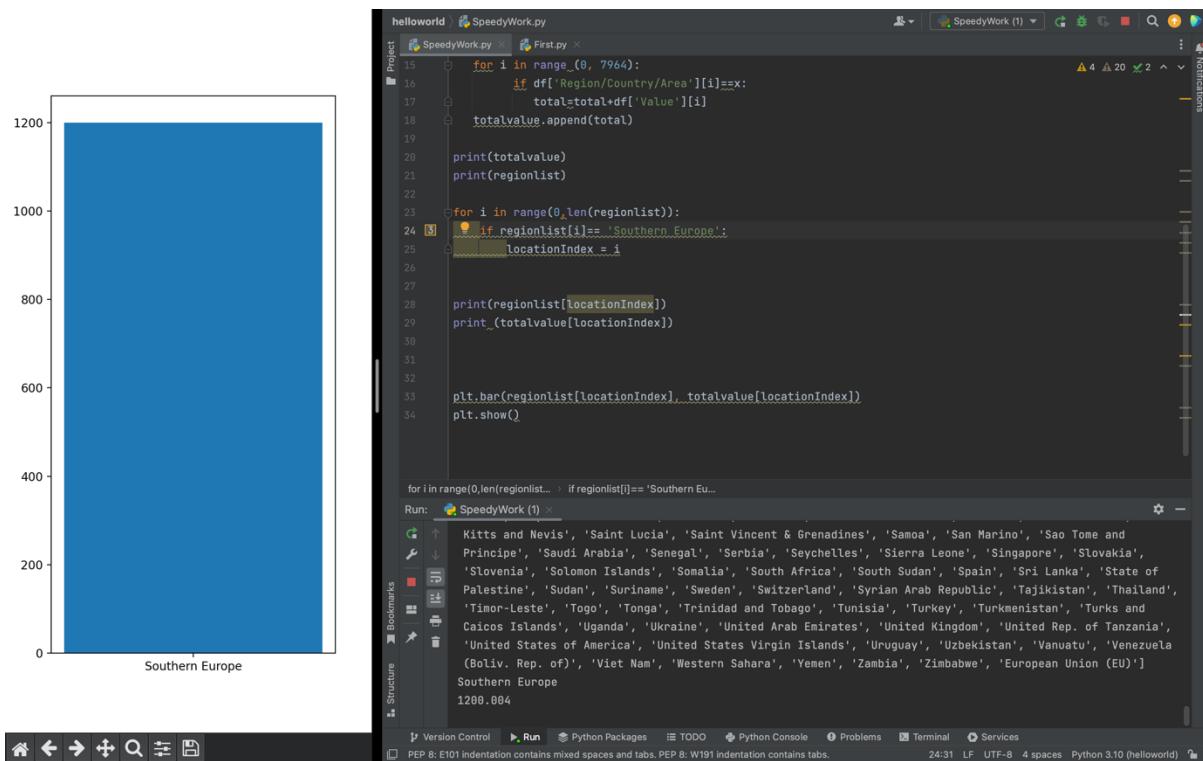
Bar chart depicting the sum of the 'Value' column per region/country/area, across all years. In this case being the Caucasus'.



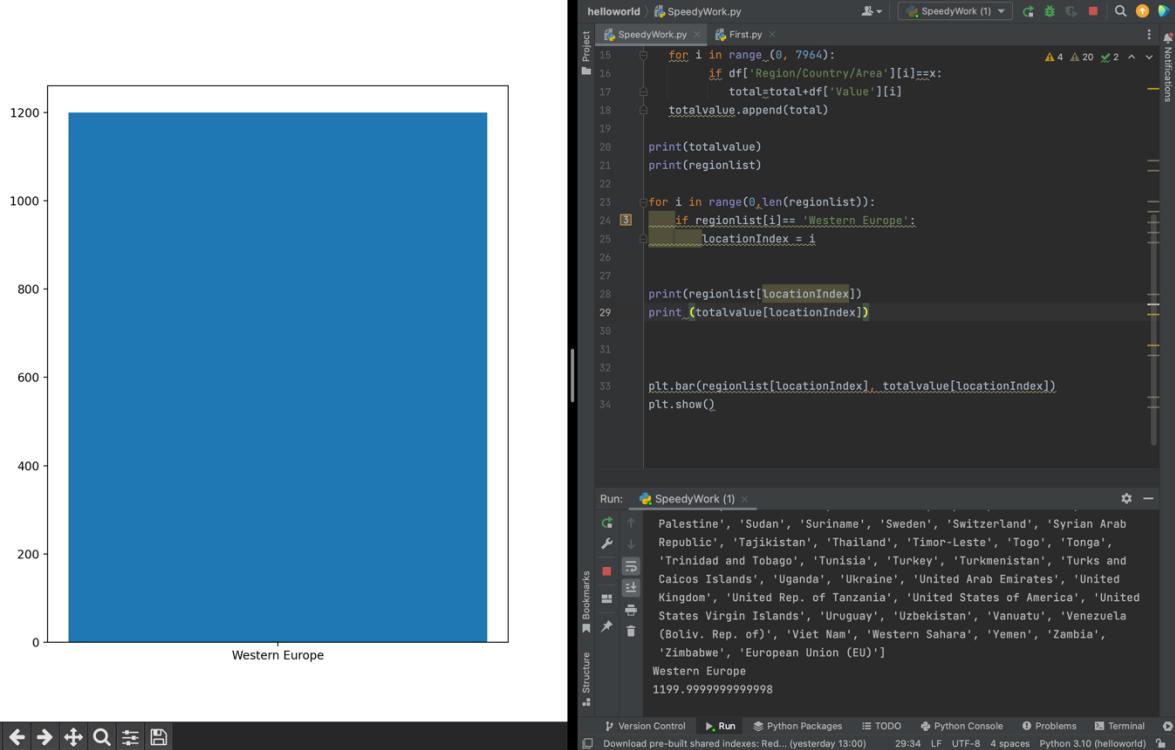
Bar chart depicting the sum of the 'Value' column per region/country/area, across all years. In this case being Eastern Europe.



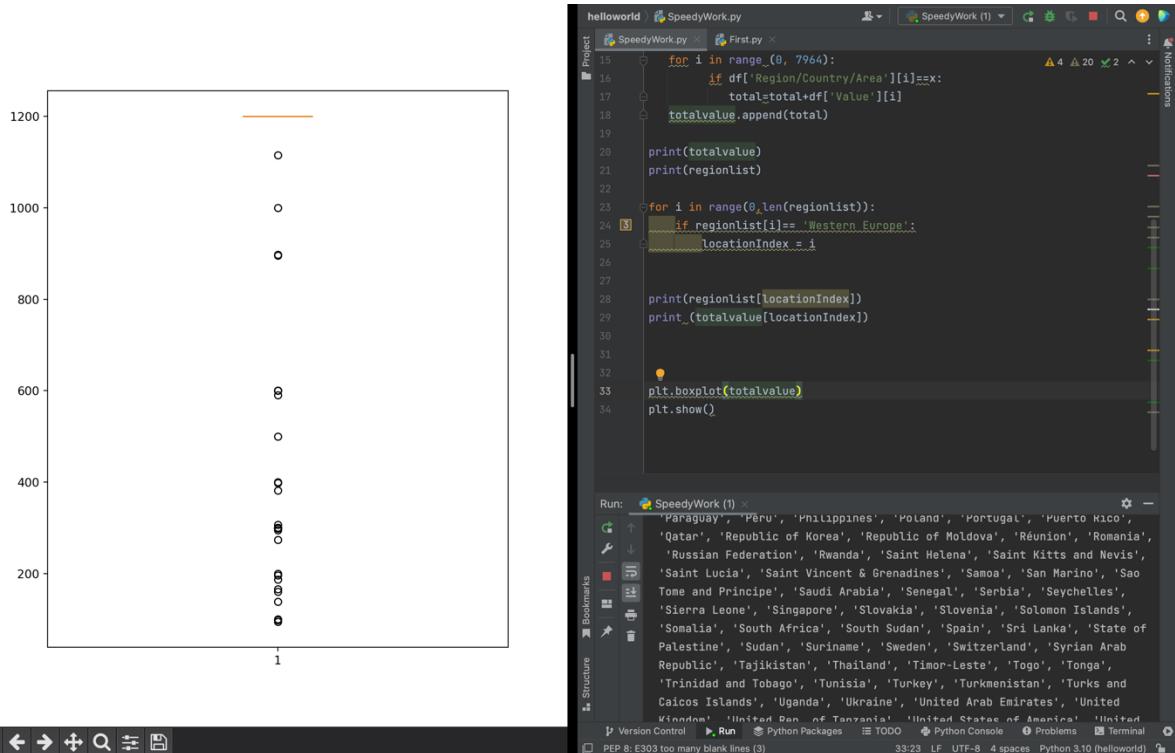
Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Northern Europe.



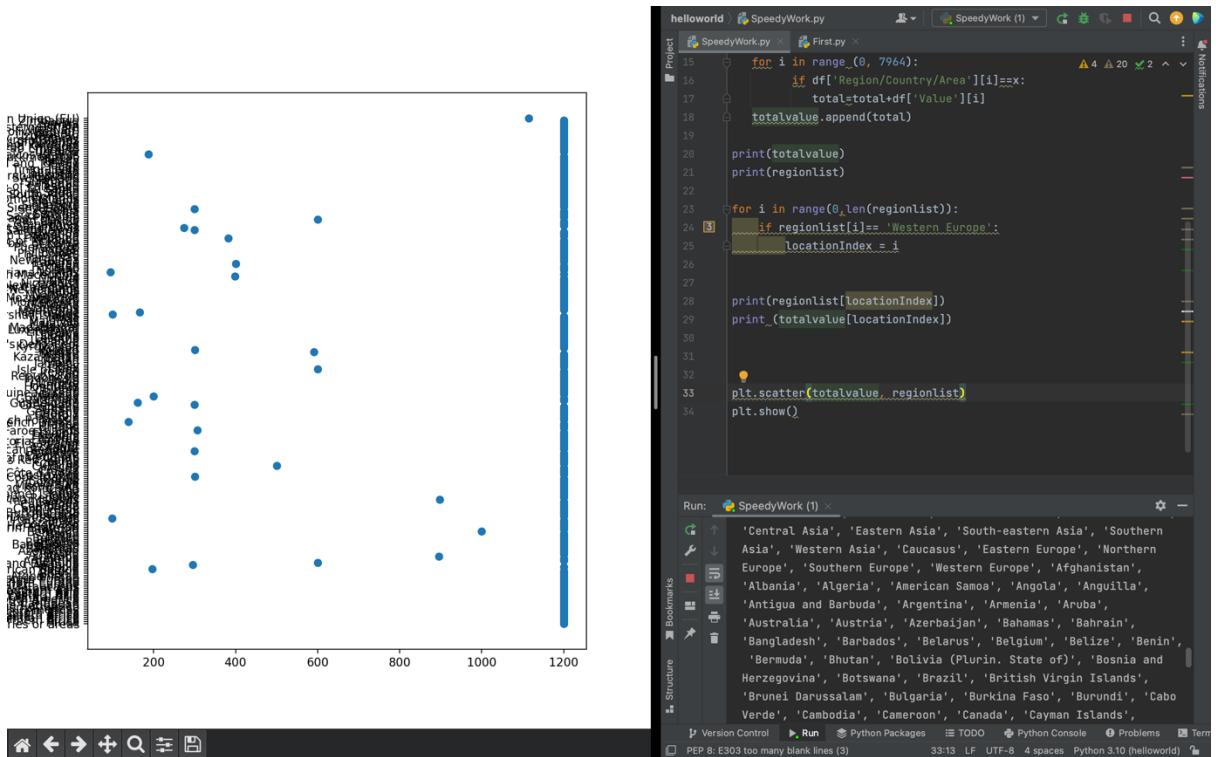
Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Southern Europe.



Bar chart depicting the sum of the 'Value' column per region/country/area, across all years. In this case being Western Europe.



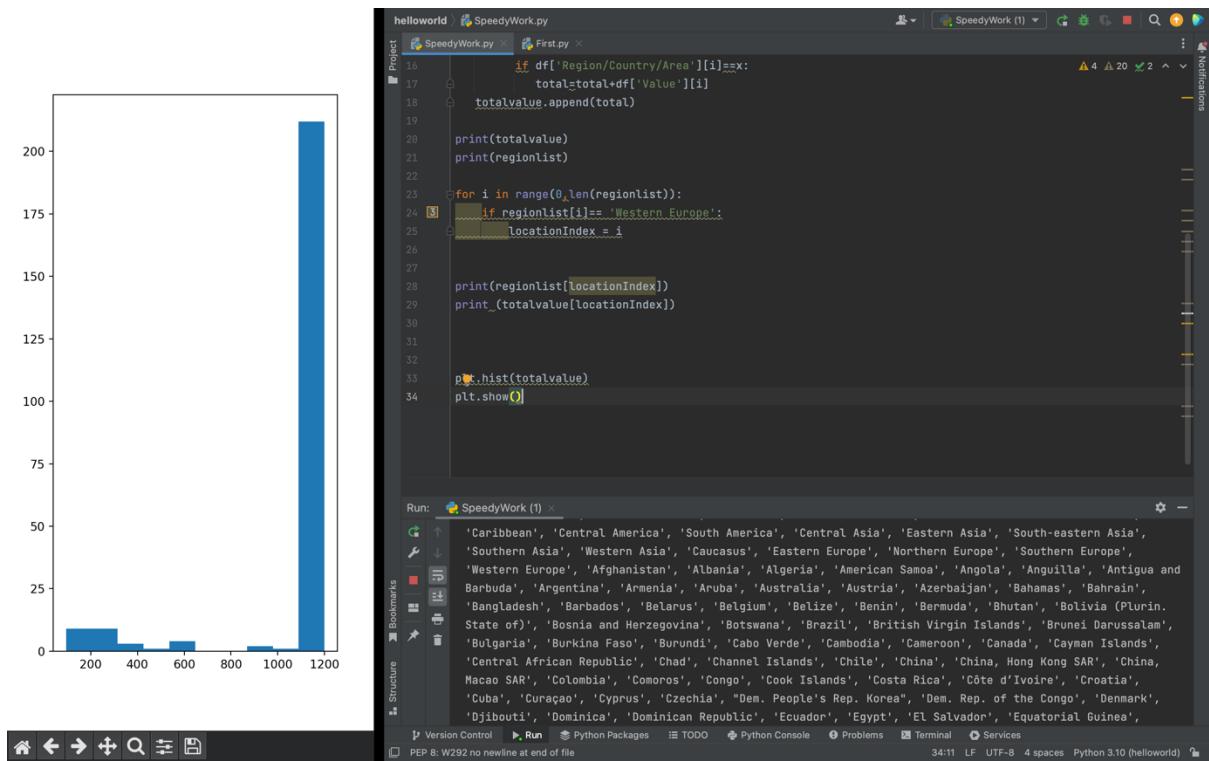
Boxplot displaying the sum of the 'Value' column. The boxplot is thin and bare due to an overwhelming majority of the data being approximately 1200.00(2dp). Every other data point below this approximation will result in an outlier being formed.



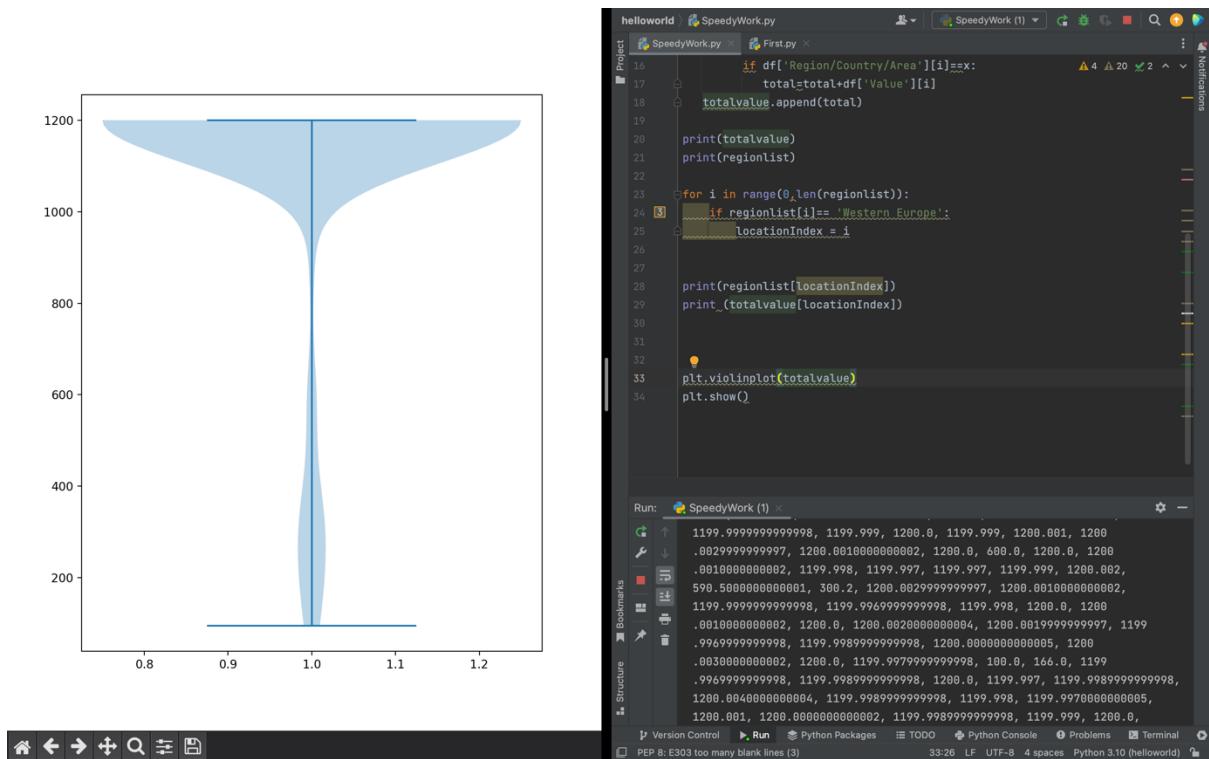
Scatter graph showing the sum of the 'Value' column. Again, this highlights the disparity in the results and where most results lay.



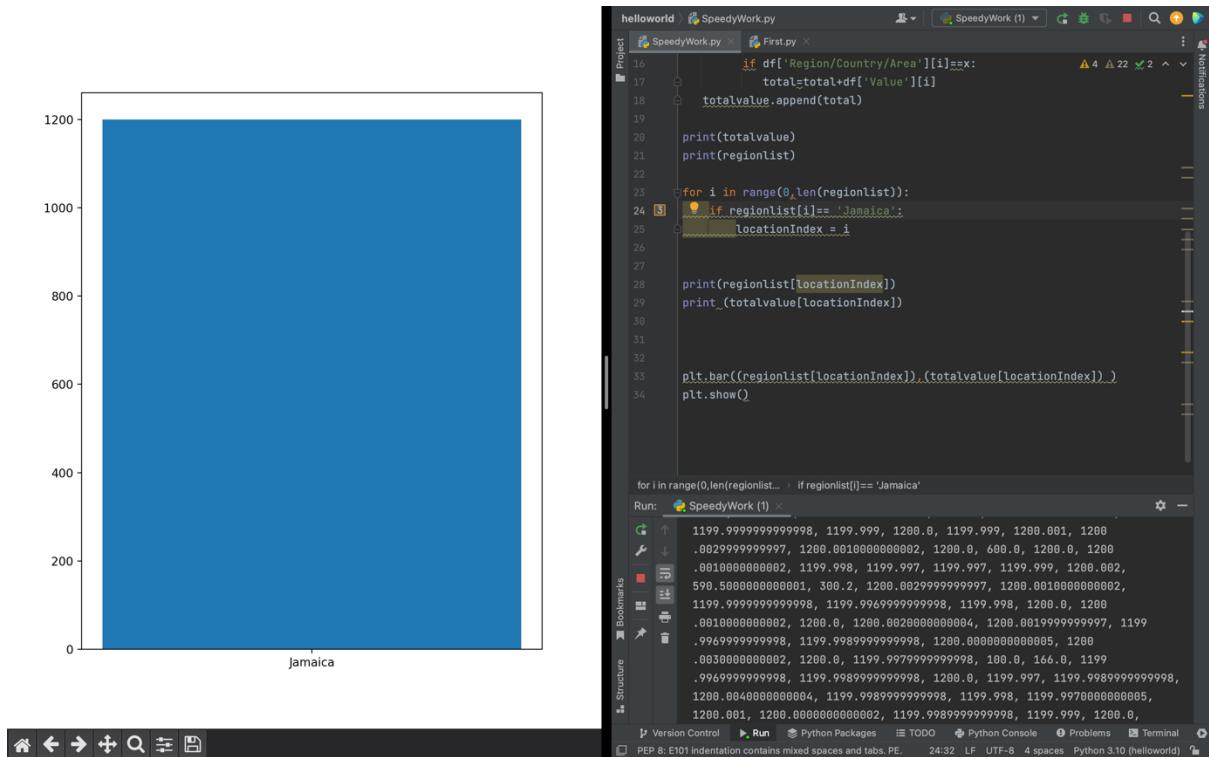
Stem graph depicting the sum of the 'Value' column. Again, this highlights the disparity in the results and where most results lay.



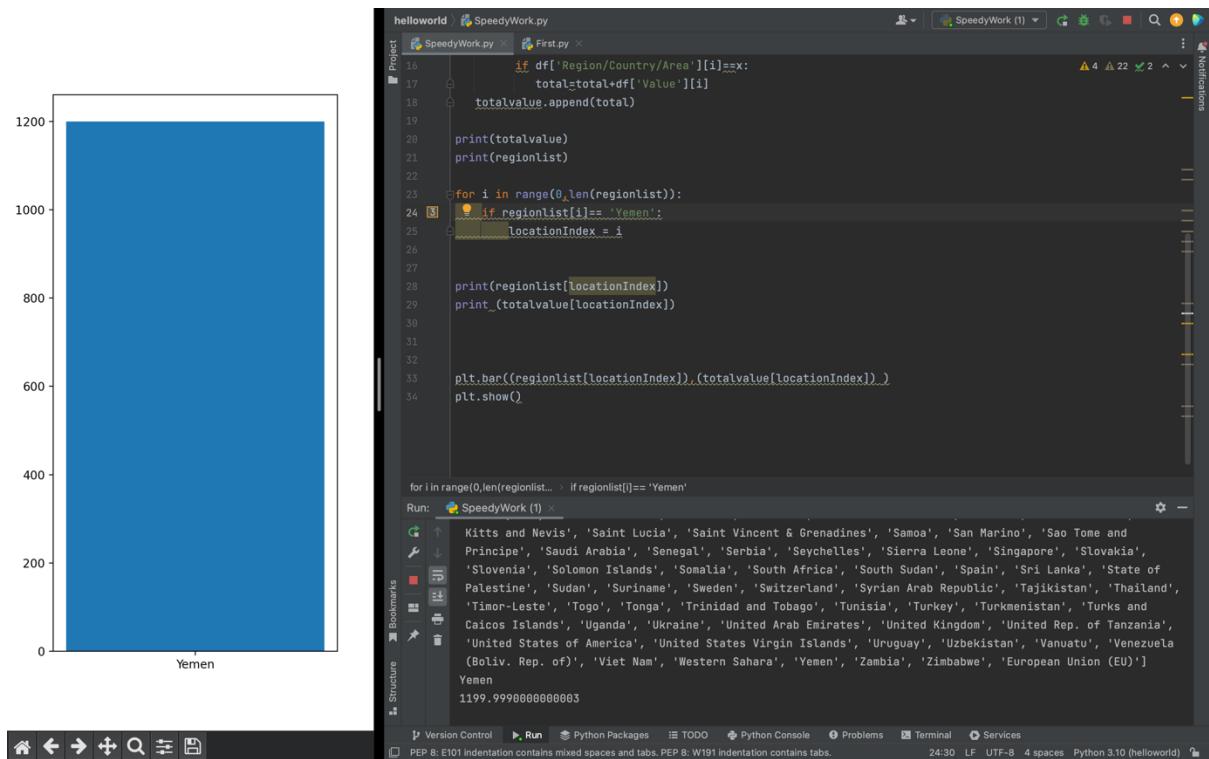
Histogram of the sum of Values in the data. Histogram does not form properly due to the density of data around 1200.00



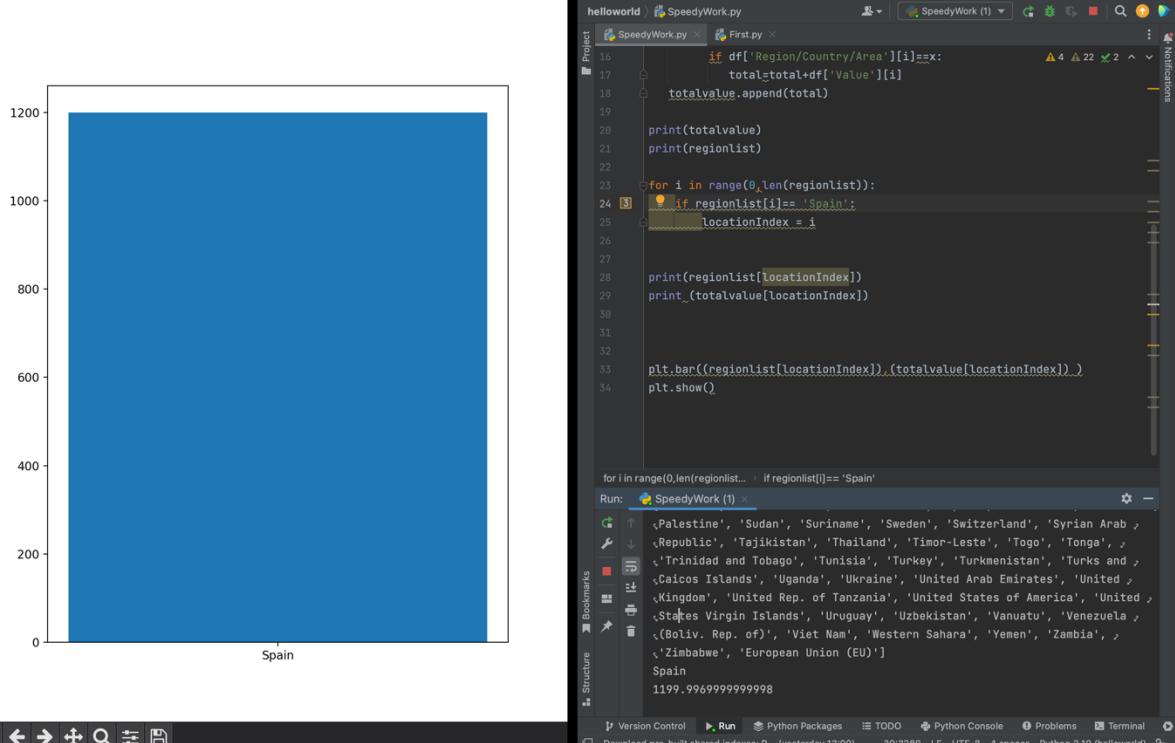
Violin plot showing the sum of the 'Value' column. Again, this highlights the disparity in the results and where most results lay. Little to no results between 700 to 900.



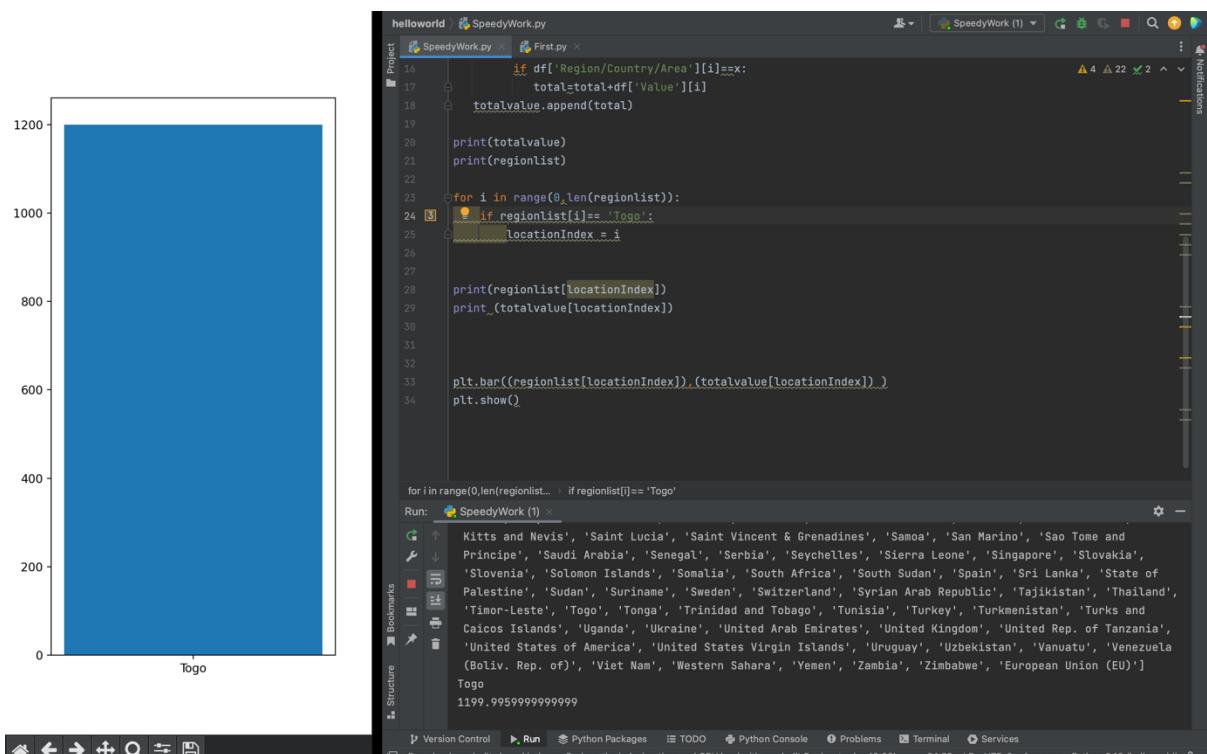
Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Jamaica.



Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Yemen.



Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Spain.



Bar chart depicting the sum of the 'Value' column per region/country/area, across all years.
In this case being Togo.