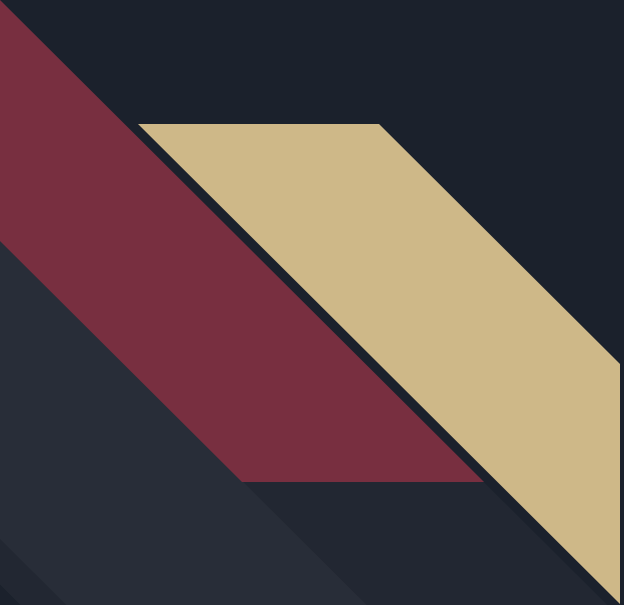




Intermediate Python Workshop Spring 2024

Jose Hernandez
Marcelina Nagales



Agenda

Goal

To run a multiple file code that can read csv files and plot data using Spyder on High Performance Compute Cluster

1

**Project Intro &
File Organization**

2

**Intro to OOD &
Spyder**

3

**Main Python
File Explanation**

4

**Function
Creation**

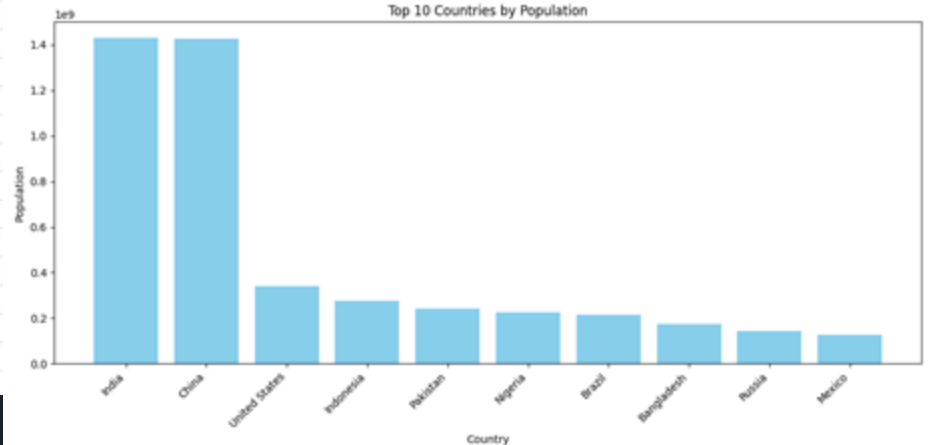
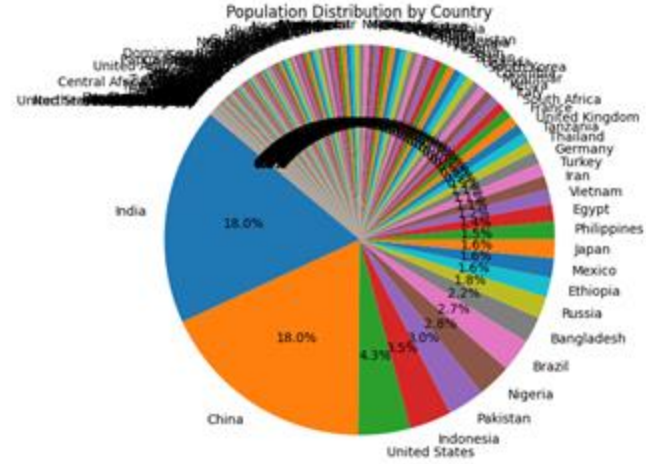
Overall project

Input a CSV of world demographics by country

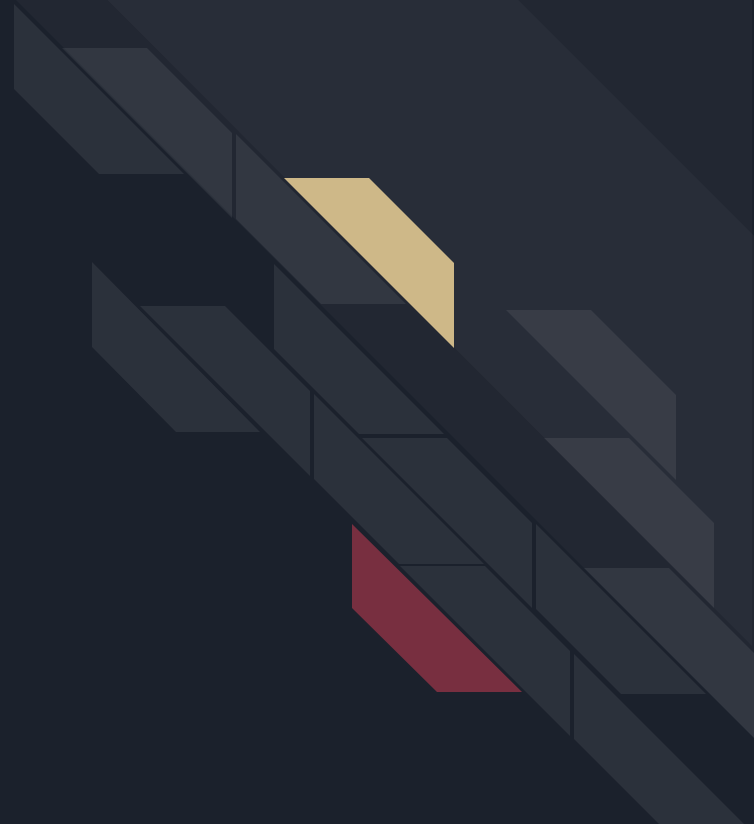
Produce 2 plots:

- A pie chart of all countries with their share of global population
- A top 10 most populous countries bar chart

1	iso_code
2	IND
3	CHN
4	USA
5	IDN
6	PAK
7	NGA
8	BRA
9	BGD
10	RUS
11	ETH
12	MEX
13	JPN
14	PHL
15	EGY
16	VNM
17	IRN
18	TUR
19	DEU
20	THA
21	TZA
22	GBR
23	FRA
24	ZAF
25	ITA
26	KEN



Setting up Spyder



Steps of Project

Input CSV

Extract Columns

Save data to New CSV

Clean said data in the new CSV

Plot the data

1	iso_code	country	2024_last_updat	2023_population	area_sq_ki	land_area_	density_/sq	growth_rat	world_%	rank
2	IND	India	1,437,330,602	1,428,627,663	3.3M	3M	485	0.92%	18.01%	1
3	CHN	China	1,425,391,096	1,425,671,352	9.7M	9.4M	151	-0.03%	17.80%	2
4	USA	United Stat	341,205,217	339,996,563	9.4M	9.1M	37	0.53%	4.27%	3
5	IDN	Indonesia	279,043,747	277,534,122	1.9M	1.9M	149	0.82%	3.50%	4
6	PAK	Pakistan	243,610,523	240,485,658	881.9K	770.9K	318	1.96%	3.06%	5
7	NGA	Nigeria	227,328,875	223,804,632	923.8K	910.8K	252	2.39%	2.86%	6
8	BRA	Brazil	217,235,074	216,422,446	8.5M	8.4M	26	0.56%	2.72%	7
9	BGD	Banglades	174,114,413	172,954,319	147.6K	130.2K	1,342	1.01%	2.18%	8
10	RUS	Russia	144,118,182	144,444,359	17.1M	16.4M	9	-0.34%	1.80%	9
11	ETH	Ethiopia	128,635,743	126,527,060	1.1M	1.1M	115	2.52%	1.62%	10
12	MEX	Mexico	129,076,931	128,455,567	2M	1.9M	67	0.73%	1.62%	11
13	JPN	Japan	122,856,549	123,294,513	377.9K	364.5K	336	-0.54%	1.53%	12
14	PHL	Philippines	118,508,757	117,337,368	342.4K	298.2K	399	1.51%	1.49%	13
15	EGY	Egypt	113,877,183	112,716,598	1M	1M	115	1.57%	1.43%	14
16	VNM	Vietnam	99,287,500	98,858,950	331.2K	313.4K	317	0.65%	1.24%	16
17	IRN	Iran	89,600,667	89,172,767	1.6M	1.6M	55	0.71%	1.12%	17
18	TUR	Turkey	86,112,117	85,816,199	783.6K	783.6K	110	0.52%	1.08%	18
19	DEU	Germany	83,268,478	83,294,633	357.1K	349.4K	238	-0.05%	1.04%	19
20	THA	Thailand	71,861,260	71,801,279	513.1K	510.9K	141	0.12%	0.90%	20
21	TZA	Tanzania	68,743,722	67,438,106	945.1K	885.8K	78	2.94%	0.87%	21
22	GBR	United King	67,885,689	67,736,802	242.9K	241.9K	281	0.33%	0.85%	22
23	FRA	France	64,840,221	64,756,584	551.7K	547.6K	118	0.19%	0.81%	23
24	ZAF	South Afric	60,795,104	60,414,495	1.2M	1.2M	50	1%	0.76%	24
25	ITA	Italy	58,757,491	58,870,762	301.3K	295.7K	198	-0.29%	0.73%	25
26	KEN	Kenya	55,826,412	55,100,586	580.4K	569.1K	99	2%	0.70%	26



File Structure

Workshop/

|— Main.py

Overall program

|— csvfunctions.py
extraction

Functions dealing with data input and

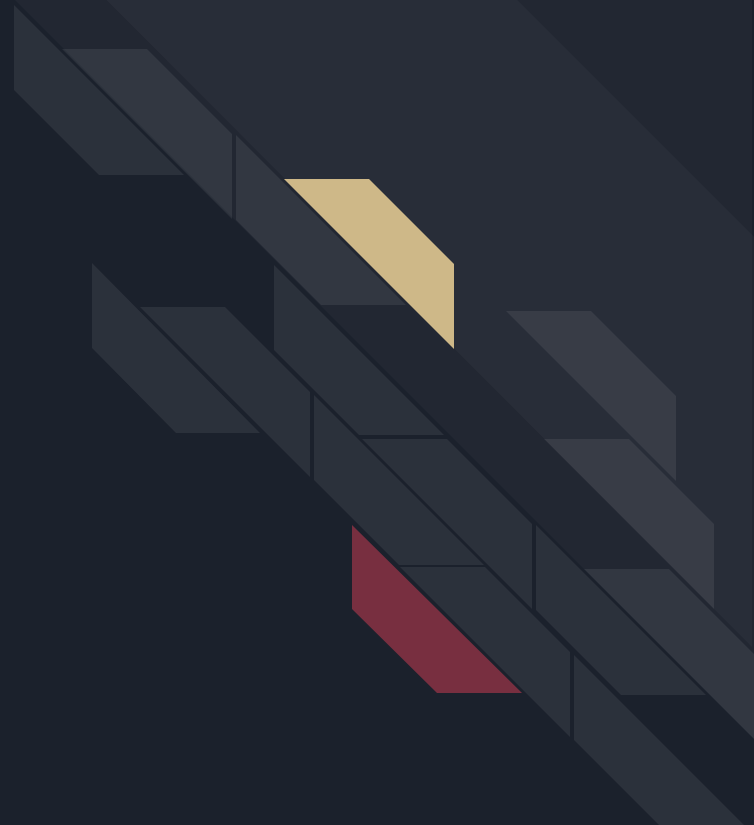
|— plotfunctions.py

Functions producing the plots

|— 2024_population.csv

Original CSV file

Wrap Up





File Structure

Workshop/

|— data/
Raw/Processed

can be split further into

|— 2024_population.csv

|— scripts/
from main.py

separate function python files

|— Main.py

|— src/
(can be project name)

function python files

|— csvfunctions.py

|— plotfunctions.py



Survey



fla.st/SKL2YUI9

We appreciate your
feedback!



Upcoming Workshops

Thank you all for attending!

Join us for other workshops hosted by the RCC in Spring 2024

- High Performance Computing (HPC) Scheduler Workshop
 - Tues April 9 at 3pm in STB 1301-E (here!)

Questions?

