IRIS 2022 Seismology Skill Building Workshop OSL

Started on	Sunday, September 11, 2022, 7:40 AM
State	Finished
Completed on	Monday, September 12, 2022, 7:56 AM
Time taken	1 day
Marks	23.67/31.00
Grade	**76.34** out of 100.00

Question 1

Correct

1.00 points out of 1.00

Flag question

In this tutorial, we will explore ways to estimate the quantity of recorded seismic noise and how it relates to human activities. In particular, we will focus on the work published in the Science paper "Global quieting of high-frequency seismic noise due to COVID-19 pandemic lockdown measures" in 2020. As the final assignment in our tutorial series, this will be the closest to actual seismology research processing we will get, which means there will be some advanced code, and a step in the processing that could take more than an hour to run - good things come to those with patience!

The study in Science calculated the RMS (root mean square) of displacement at nearly 200 seismic stations around the world. The primary way that the authors have done this is by estimating the power spectral density (PSD) of seismic noise recorded over days of time at each station. The PSD is just a statistical way of representing the seismic energy (power) at different frequencies (spectral), by estimating the number of times that amount of energy at that frequency is observed over a period of time (density). I found this short article about PSDs to be a nice introduction to help folks understand them:

https://vru.vibrationresearch.com/lesson/what-is-the-psd/

Why is the power an important quantity to measure instead of average amplitude?

Select one:

- a. It removes the instrument response
- b. It takes into account the physical units
- c. It takes into account the absolute value of vibrations around a mean of zero ✓
- d. It estimates the energy involved

Check

Correct
Marks for this submission: 1.00/1.00.

Question 2

Correct

1.00 points out of 1.00

Flag question

The lead author of the Science article on the reduction of seismic noise during COVID-19 is Thomas Lecocq (from Belgium), so you can find the Jupyter notebook used by authors in the study by looking him up on GitHub. On his GitHub page, you should find a repository for his shared Jupyter notebook for calculating the RMS of a seismic signal from PSDs. What is the name of this repository?

Select one:

- a. msnoise-sara
- b. SeismoRMS ✓
- c. MSNoise
- d. 2020_Science_GlobalQuieting
- e. msnoise-tomo

Check

Correct
Marks for this submission: 1.00/1.00.

Question 3

Correct

1.00 points out of 1.00

Flag question

Now log into OSL and move into the **jupyter** directory you created in a previous assignment. Once you are inside the **jupyter** directory, which command do you type to download the GitHub repository to this directory?

Answer: git clone "https://github.com/ThomasLecocq/SeismoRMS.git" ✓

Check

Correct. Go ahead and run this now: `git clone "https://github.com/ThomasLecocq/SeismoRMS"`

Correct
Marks for this submission: 1.00/1.00.

Question 4

Correct

1.00 points out of 1.00

Flag question

When you run this command it should create a directory with the name of the repository and then download a set of files into that directory. Which of the files is a Jupyter notebook we can use for this tutorial?

Select one:

- a. README.md
- b. seismosocialdistancing.py
- c. environment.yml
- d. SeismoRMS.py
- e. SeismoRMS.ipynp
- f. SeismoSocialDistancing.ipynb ✓

Check

Correct
Marks for this submission: 1.00/1.00.

Question 5

Correct

0.67 points out of 1.00

Flag question

Go ahead and run `jupyter notebook` from the `SeismoRMS` directory and then open the Jupyter notebook file. When the Jupyter Notebook loads, take a minute to scroll through the notebook to see what it is doing. What is the order of things that are done by this notebook?

Import libraries	1 ✓
Calculate the root mean square motion	4 ✓
Create displacement plots	6 ✓
Get seismic data	2 ✓
Create amplitude plots	5 ✓
Compute power spectral density	3 ✓

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.67/1.00**.

Question 6

Correct

1.00 points out of 1.00

Flag question

The first Markdown cell at the beginning of the notebook should provide a brief introduction that lists the library requirements for the code to run properly. We have sought to ensure your OSL environment has this software installed already, so there is no need to set up the conda environment on your own.

There are plenty of libraries you will have seen before, but there are some new functions and libraries involved in this code, so we should take a minute to learn about them. For example, what does the`from glob import glob`command do?

Select one:

- a. load a library for finding system files, primarily using wildcards ✓
- b. load a function for searching for text in strings
- c. load file system functions
- d. load a library for combining files together

Check

Correct
Marks for this submission: 1.00/1.00.

Question 7

Correct

1.00 points out of 1.00

Flag question

What does the tqdm library do?

Select one:

- a. sorts a list of dates
- b. creates a progress bar output for loops ✓
- c. tracks progress of downloads
- d. converts time and date

Check

Correct
Marks for this submission: 1.00/1.00.

Question 8

Correct

1.00 points out of 1.00

Flag question

What does the `import seismosocialdistancing` command at the end of this Code cell do?

Select one:

- a. loads the seismosocialdistancing.ipynb Jupyter notebook
- b. loads the seismosocialdistancing function from Python core
- c. loads the seismosocialdistancing.py Python script ✓
- d. loads the seismosocialdistancing library from ObsPy

Check

Correct
Marks for this submission: 1.00/1.00.

Question 9

Correct

0.00 points out of 1.00

Flag question

Since the lead author Thomas Lecocq works at the Royal Observatory of Belgium, the default station for the Jupyter notebook is in Brussels, Belgium, which is obtained from the ODC data provider. Since we are based in the United States, I am going to recommend that we choose a station in the US that we can obtain from the IRIS DMC. Considering the severe outbreak in the New York, I will recommend that we change the parameters to those for station CPNY in Central Park, New York. Take a minute to use one of the many tools we have learned in this course to find out the rest of the information for this station, assuming we want data from a broadband, vertical channel, recording at 40 samples per second.

location	"" ✓
channel	BHZ ✓
network	LD ✓

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.00/1.00**.

Question 10

Correct

0.67 points out of 1.00

Flag question

Be sure to change the station information in your code if you haven't already. We also need to change the Time zone (America/New_York), Site description (Central Park, NY), and Data provider (IRIS). We should also set the time of the bans/closures put in place in New York City, which happened on March 16 for Restaurants/Bars/Schools, and March 23 for Non-essential shops. Go ahead and make these changes to the code.

We can keep the default start and end dates for the seismogram request even though there is more data to process. This will make sure it does not take too long to download and process. What are the start and end dates based on the default code settings?

Select one or more:

- a. Start Date: January 1, 2020
- b. End Date: June 1, 2020
- c. Start Date: May 1, 2020
- d. Start Date: March 1, 2020 ✓
- e. End Date: May 1, 2020 ✓
- f. Start Date: April 1, 2020
- g. End Date: March 1, 2020
- h. End Date: April 1, 2020

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.67/1.00**.

Question 11

Correct

1.00 points out of 1.00

Flag question

Make sure you have run the Code cell in Step 1 and Step 2.

Then run the Code cell in Step 3 to download the data to your machine. This takes about 10 minutes, depending on the speed of your internet connection. In the meantime, you can take a closer look at the code. Inside the `for loop` there is a get_waveforms() call to get the data. Based on the times being input for the start and end times, how much time in minutes on either side of the beginning and end of the day is being requested?

Answer: 30.0 ✓

Check

Correct
Marks for this submission: 1.00/1.00.

Question 12

Correct

1.00 points out of 1.00

Flag question

Once the code is done processing (it may take a while!), you can use another Terminal window to go into the jupyter/SeismoRMS directory to see what has been created by the notebook. When you look inside this directory, what do you find that is different than the repository files?

Select one:

- a. A miniseed file for each day ✓
- b. A miniseed file for each hour
- c. A metadata file for each month
- d. A miniseed file for each month
- e. A metadata file for the whole timeframe
- f. A metadata file for each hour
- g. A miniseed file for the whole timeframe
- h. A metadata file for each day

Check

Correct
Marks for this submission: 1.00/1.00.

Question 13

Correct

1.00 points out of 1.00

Flag question

Then run the Code cell in Step 4 to compute the PSDs. This step takes close to an hour to run - sorry for the inconvenience of this, but this is often what it takes to process seismic data for research purposes.

While the Step 4 code is running, we can peek at the code to see what is happening. This code is more complex than most we have looked at, but hopefully you can still pick out some interesting details.

In order to convert the waveforms to displacement, the code needs to use the instrument response. Looking at this Code cell and the one before it, which function obtains the instrument response from the data center?

Select one:

- a. PPSD()
- b. read()
- c. attach_response()
- d. get_waveforms() ✓

e. get_stations()

Check

Correct
Marks for this submission: 1.00/1.00.

Question 14

Correct

0.00 points out of 1.00

Flag question

The heart of this code is the PPSD() function that calculates the Probabilistic Power Spectral Density. You can read more about this function and how to use it at these two sites:

https://docs.obspy.org/packages/autogen/obspy.signal.spectral_estimation.PPSD.html

https://docs.obspy.org/tutorial/code_snippets/probabilistic_power_spectral_density.html

The PPSD function is called in this code with a ppsd_length of 1800 and a 50% overlap. Based on this, how many power estimates will be made in a single day for a given frequency?

Answer: 96

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.00/1.00**.

Question 15

Correct

1.00 points out of 1.00

Flag question

The PPSD library can also create plots based on the calculated PSDs, which will be used later in this notebook. Based on the manual, when the typical PSD plot is generated, it will show a scale bar to the right of the plot that illustrates the density of the observations. What does the scale bar below the plot show?

Select one:

- a. The available data over time and when single PSD measurements go into the plot ✓
- b. The available of data over time and when PSD measurements are within the NLNM and NHNM
- c. The amplitude data over time and when single PSD measurements go into the plot
- d. The amplitude of data over time and when PSD measurements are within the NLNM and NHNM

Check

Correct
Marks for this submission: 1.00/1.00.

Question 16

Correct

1.00 points out of 1.00

Flag question

When the Step 4 code is done running, use the second Terminal window to check again to see what has been created.

Select one:

- a. A .mseed file for each month
- b. A .npz file for each hour
- c. A .mseed file for each day
- d. A .mseed file for each hour
- e. A .mseed file for the whole time frame
- f. A .npz file for each day ✓
- g. A .npz file for the whole time frame
- h. A .npz file for each month

Check

Correct
Marks for this submission: 1.00/1.00.

Question 17

Correct

1.00 points out of 1.00

Flag question

Then run the Code cell in Step 5 to reload the PSDs for each day. Fortunately, this should take less than a minute to complete. Which data is being read by this code?

Select one:

- a. The local .npz files ✓
- b. The local .ipynb files
- c. The local .mseed files
- d. The local .yml files
- e. The local .py files

Check

Correct
Marks for this submission: 1.00/1.00.

Question 18

Correct

0.67 points out of 1.00

Flag question

Then run the Code cell in Step 6 to create the first set of plots. You may see a pink warning generated by this code, but it just provides some advice for importing functions.

The first plot it should produce would be a traditional PSD plot with amplitude on the y-axis and Period on the x-axis. The lighter colors show the most common spectral content of this stations over the 2 months of data we analyzed. Human generated noise is generally thought to be above 1 Hz, so which side of this plot shows the observations above 1 Hz?

Select one:

- a. lower right
- b. lower left ✓
- c. upper left
- d. upper right

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.67/1.00**.

Question 19

Correct

0.33 points out of 1.00

Flag question

The second plot the code should produce would be the RMS seismic amplitude over time. This plot tends to be a little small to interpret, but which fluctuations in noise level can be seen in this plot?

Select one or more:

- a. pre-lockdown vs. post-lockdown ✓ 1 of 3 correct answers
- b. day vs. night ✓ 1 of 3 correct answers
- c. winter vs. summer
- d. weekday vs. weekend ✓ 1 of 3 correct answers
- e. holiday vs. non-holiday

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.33/1.00**.

Question 20

Correct

0.33 points out of 1.00

Flag question

The third plot should be a spectrogram of the whole time, but it can take a really long time to load. The [*] symbol to the left of the Code cell should indicate it is still running when you read this. Although you could wait for this to finish, it is not important for this tutorial, so instead it gives us a chance to discuss how to stop the kernel for a Jupyter notebook and how to start it over again. Click the Stop button just to the right of the Run button in the toolbar at the top. Then click on the circular Restart button that is just to the right of the Stop button. It will ask you if you are really sure you want to Restart, and yes confirm you want to Restart. Once you do this, the [*] symbol should turn back to [ ] to indicate the Code has not been run. Then which command do you need to comment out with a # symbol at the beginning of the line?

Select one:

- a. register_matplotlib_converters()
- b. [ppsd.plot_temporal(0.10) for mseedid, ppsd in ppsds.items()]
- c. [ppsd.plot(max_percentage=3) for mseedid, ppsd in ppsds.items()]
- d. [ppsd.plot_spectrogram(clim=(-160,-100)) for mseedid, ppsd in ppsds.items()] ✓ Please make sure to comment this line out.

e. from pandas.plotting import register_matplotlib_converters

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.33/1.00**.

Question 21

Correct

0.33 points out of 1.00

Flag question

Since you restarted the kernel, you will need to go back to the beginning of the notebook, and run each of the Code cells in order from the beginning, starting from Step 1 all the way up through the revised version of Step 6. Fortunately, the data is already downloaded and the PSDs are already calculated so each of those files will be read in much more quickly this time.

Once you have re-run each Code cell in order, you can also run the code in Step 7, which only takes about 30 seconds. Next run the second code cell labeled Weekday/Time of day Analysis. This should produce an easier to read plot of the one in Step 6.

Based on when the bans went in place and when the seismic noise actually came down to significantly lower levels, when does it look like the biggest change in human generated noise occurred?

Select one:

a. April April 13

b. Around March 23

c. Around March 16

d. Around March 30 ✓

e. Around April 6

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.33/1.00**.

Question 22

Correct

1.00 points out of 1.00

Flag question

You can change the frequency ranges in the args structure. What is the default frequency band?

Select one:

a. 0.1-1.0

b. 4.0-14.0 ✓

c. 4.0-20.0

d. 1.0-20.0

Check

Correct
Marks for this submission: 1.00/1.00.

Question 23

Correct

1.00 points out of 1.00

Flag question

I have indicated that human generated noise is most pronounced in the higher frequencies, but we can test this by changing the frequency band to the lowest frequency range from the options in the previous question. What does the plot look like when you run the code again with this frequency band?

Select one:

a. A clear decrease in overall noise level after the lockdowns began

b. A slight decrease in overall noise level after the lockdowns began, but several odd spikes in noise level too

c. A slight increase in overall noise level after the lockdowns began, but several odd spikes in noise level too

d. Variable but no clear change in seismic noise behavior during the lockdowns ✓

e. A clear increase in overall noise level after the lockdowns began

Check

Correct
Marks for this submission: 1.00/1.00.

Question 24

Correct

0.67 points out of 1.00

Flag question

We can also try the widest range of frequencies to see if that improves the plot. What does the plot look like when you run the code again with this frequency band?

Select one:

a. A slight decrease in overall noise level after the lockdowns began, but several odd spikes in noise level too ✓

b. Variable but no clear change in seismic noise behavior during the lockdowns

c. A slight increase in overall noise level after the lockdowns began, but several odd spikes in noise level too

d. A clear decrease in overall noise level after the lockdowns began

e. A clear increase in overall noise level after the lockdowns began

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.67/1.00**.

Question 25

Correct

0.33 points out of 1.00

Flag question

Return to the 4.0-14.0 frequency band and remake the plot. What is the average peak weekday displacement for the first week of March in nanometers? AKA - around what value do the peaks cluster around?

Answer: 2.3 ✓

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.33/1.00**.

Question 26

Correct

1.00 points out of 1.00

Flag question

What is the average peak weekday displacement in nanometers during the quarantine in April?

Answer: 1.5 ✓

Check

Correct
Marks for this submission: 1.00/1.00.

Question 27

Correct

1.00 points out of 1.00

Flag question

What is the reduction in percent of peak weekday seismic noise between March and April?

Answer: 34.7 ✓

Check

Correct
Marks for this submission: 1.00/1.00.

Question 28

Correct

1.00 points out of 1.00

Flag question

Go ahead and run the next Code cell that would produce the 'dayplots' type. The solid lines are for pre-lockdown, the dashed lines are for post-lockdown, and the colors show the day of the week (see legend). Which day of the week during pre-lockdown has the most similar curve to the noise pattern of a weekday during the lockdown?

Select one:

a. Tuesday

b. Thursday

c. Friday

d. Saturday ✓ Yes, a Wednesday during lockdown is like a normal Saturday.

e. Sunday

f. Monday

g. Wednesday

Check

Correct
Marks for this submission: 1.00/1.00.

Question 29

Correct

1.00 points out of 1.00

Flag question

Move on to run the next Code cell that would produce the 'clockplots' type. These plots show hours of the day going around in a circle, with the displacement plotted as distance from the center of the circle. Approximately what hour of the day (local time) is the seismic noise largest in Central Park during the week?

Answer: 8

Check

Correct
Marks for this submission: 1.00/1.00.

Question 30

Correct

0.33 points out of 1.00

Flag question

Move on to run the next Code cell that would produce the 'clockmaps' type. These plots show displacement on the color scale (darker is quieter), with hours of the day going around in a circle, and the days of the calendar plotted as distance from the center of the circle. Time goes from the inside out, starting at the beginning of March in the center of the circle and ending at the beginning of May on the outer rim. The closing of schools and stores can be seen as colored circles. What are the series of about 9 darker concentric circles seen in this plot?

Select one:

- a. day time
- b. the lockdown
- c. weekdays
- d. night time
- e. weekends ✓

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.33/1.00**.

Question 31

Correct

0.33 points out of 1.00

Flag question

Move on to run the next Code cell that would produce the 'gridmaps' type. These plots show displacement on the same color scale (darker is quieter), with hours of the day going from bottom to top, and the days of the calendar plotted from left to right. Is it subtle, but I think you can pick out which week when there is a noticeable increase in seismic activity compared to the prior week. Still, it may help to return to the first plot in this section that showed the displacement over time, and note that there is a orange line that shows the time-averaged values for hours 6-16 of the day. So which week appears to be the first one where the activity starts increasing again?

Select one:

- a. Week of April 8
- b. Week of May 6
- c. Week of April 29
- d. Week of April 15
- e. Week of April 22 ✓

Check

Correct
Marks for this submission: 1.00/1.00. Accounting for previous tries, this gives **0.33/1.00**.

Finish review