**Outlook**

---

**Meeting Summary for Course: Geocomputation and geospatial analysis**

---

**From** Meeting Summary with AI Companion <no-reply@zoom.us>
**Date**  Tue 4/1/2025 12:18 PM
**To**       Amatulli, Giuseppe <giuseppe.amatulli@yale.edu>

**NEW**  Meeting summary with AI Companion now supports additional languages in preview.                    **Learn More**

## Meeting summary for Course: Geocomputation and geospatial analysis (04/01/2025)

### Quick recap

Giuseppe introduced the course on geocomputation and geospatial analysis, focusing on bash language, Linux operating systems, and Python, with an emphasis on language interoperability in the Linux environment. He also explained the basics of using the Bash terminal, the difference between text files and binary files, and file naming conventions in Linux. Additionally, he demonstrated various bash commands for working with large text files, introduced Jupyter Lab as a tool for the next class, and addressed troubleshooting virtual machine installation issues.

### Next steps

• All attendees to install and set up the virtual machine before Thursday's class.
• All attendees to practice using bash commands and become familiar with text file manipulation.
• All attendees to prepare their own text or CSV files for practice during the next class.
• Saverio to provide a link to Linux keyboard shortcuts for the attendees.
• Giuseppe to post the link for today's presentation PDF, video recording, and awk video recording on the syllabus page.
• Sofia to convert her Excel files to text or CSV format for use in bash.
• Saverio to investigate potential workarounds for students with ARM processors under Windows having issues with VirtualBox.
• All attendees to review the awk tutorial video that Giuseppe will provide.
• Saverio to research existing web GIS systems for automatic landslide detection and mapping, focusing on operational and dynamic solutions.

## Summary

### Geocomputation Course Overview and Setup

Giuseppe introduces the course on geocomputation and geospatial analysis, organized by the Italian Integrated Environmental Research Infrastructure System. He explains that the course will focus on bash language, Linux operating systems, and Python, with an emphasis on language interoperability in the Linux environment. Giuseppe highlights the advantages of using Linux for geospatial analysis, including its versatility, low hardware requirements, and ability to integrate multiple programming languages. He introduces the virtual machine that students will use for the course and demonstrates how to set it up. Giuseppe emphasizes the importance of staying within the virtual machine environment for consistency and encourages students to explore the various software tools available. The class will begin with basic bash commands before moving on to GDAL and Python.

### Bash Terminal Basics and File Handling

Giuseppe explains the basics of using the Bash terminal, including command syntax, options, and file redirection. He demonstrates how to list directory contents, search for commands, and count words in a file. Giuseppe emphasizes the importance of understanding user and computer names in the terminal prompt for file transfers. He shows how to redirect standard output and errors to files, and explains the difference between overwriting and appending to files. Saverio adds that many sensors use Bash for logging data and errors, highlighting the practical application of these concepts in data collection and error reporting.

### File Types and Command Line Operations

Giuseppe explains the difference between text files and binary files, emphasizing that file extensions don't always accurately represent the file type. He demonstrates how to use special characters like asterisks and question marks for file listing and searching. Giuseppe also introduces the concept of pipes in command-line operations, showing how to chain commands together to perform more complex tasks. He compares this functionality to similar concepts in programming languages like R and Python.

### Piping in Data Manipulation With Awk

Giuseppe explained the concept of piping in data manipulation, using the output of one command as input for another. He demonstrated this with examples, including using the 'awk' command for text manipulation. Giuseppe emphasized the power and speed of 'awk' for tasks like selecting specific columns in a text file. He also encouraged the use of piping in future scripts.

### Linux File Naming and Navigation

Giuseppe explains file naming conventions in Linux, advising against using spaces in filenames and recommending underscores instead. He demonstrates how to navigate directories using the command line, emphasizing the importance of using the tab key for auto-completion to avoid errors and increase efficiency. Giuseppe also covers basic Linux commands like 'cd', 'ls', and discusses file paths, including the use of the tilde symbol. He introduces the concept of working with large text files in Bash, explaining how Bash

handles files differently from programs like Excel by keeping the file on disk and only loading portions into RAM as needed.

**Bash Commands for Large Text Files**
Giuseppe demonstrates various bash commands for working with large text files containing fire data. He shows how to count lines, sort data numerically and alphabetically using the 'sort' command with different options, and filter data using 'grep'. Giuseppe explains how to use 'awk' for summing values across multiple files efficiently. He then covers looping through files and years using for loops, demonstrating different syntax options. Throughout, Giuseppe emphasizes the speed and efficiency of bash commands for processing large datasets compared to other programming languages.

**Text File Analysis and Tools**
Giuseppe discussed the use of text files in data analysis, emphasizing the importance of specifying delimiters when working with comma-separated values (CSV) or other non-default formats. He encouraged team members to share their text files for practice and suggested using tools like Git Copilot for assistance. Saverio added that while Bash and Awk are efficient for large data sets, Python with Pandas is better suited for structured data. The team was advised to explore their data and consider the appropriate tool based on the data's complexity and structure.

**Jupyter Lab Installation and Usage**
Giuseppe introduces Jupyter Lab as a tool for the next class and demonstrates how to install and use it. He explains that Jupyter Lab allows for both Markdown and code cells, with code cells defaulting to Python unless specified otherwise. To run Bash commands in Jupyter Lab, an exclamation mark or '%%bash' at the beginning of a cell is needed. Giuseppe encourages students to start using Jupyter Lab with their data, showing how to navigate directories and access files. He concludes by mentioning that the next class will focus on geographical data and reminds students to practice with Bash commands.

**Troubleshooting Virtual Machine Installation Issues**
Giuseppe and the team discuss troubleshooting virtual machine installation issues with several participants. Enrica successfully follows the installation steps with Giuseppe's guidance, including mounting directories and rebooting. Sofia and others are advised to follow similar steps if they encounter problems. Saverio warns about potential issues with ARM processors on Windows machines. The meeting concludes with a reminder that everyone needs to have their virtual machines ready by Thursday, and assistance will be available the following day for those still facing problems.

*AI-generated content may be inaccurate or misleading. Always check for accuracy.*

Please rate the accuracy of this summary.  👍  👎
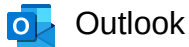
[ Share summary ]   **Edit**

 Best,

Zoom

+1.888.799.9666

©2025 Zoom Communications, Inc.

Visit zoom.com

55 Almaden Blvd
San Jose, CA 95113

**Outlook**

---

**Riepilogo riunione per Course: Geocomputation and geospatial analysis**

---

**From** Meeting Summary with AI Companion <no-reply@zoom.us>
**Date** Wed 4/2/2025 10:37 AM
**To**      Amatulli, Giuseppe <giuseppe.amatulli@yale.edu>

---

**NEW**   Riepilogo della riunione con AI Companion ora supporta ulteriori lingue in anteprima.    **Ulteriori informazioni**

## Riepilogo della riunione per Course: Geocomputation and geospatial analysis (04/02/2025)

### Riepilogo rapido

Il riassunto della riunione menziona che Luigi ha fatto riferimento a una data specifica, probabilmente il secondo o terzo giorno di qualcosa, senza fornire ulteriori dettagli sul contesto. Luigi ha ripetuto più volte l'espressione "ogni giorno" durante la discussione. La riunione si è conclusa con Luigi che ha pronunciato "Od Hood", una frase il cui significato non è chiaro dal contesto fornito.

### Fasi successive

Next steps were not generated due to insufficient transcript.

---

### Riepilogo

**Scusa, Ma Non C'è Un Riassunto Di Incontro Fornito. Potresti Per Favore Fornire Il Riassunto Dell'incontro in Modo Che Possa Generare Un Titolo Di Incontro Appropriato?**
Luigi menziona che oggi è il giorno 2 o 3, ma non è chiaro a cosa si riferisca specificamente. Egli ripete "ogni giorno" e conclude con "Od Hood", il cui significato non è evidente dal contesto fornito.

*Il contenuto generato da intelligenza artificiale potrebbe essere impreciso o fuorviante. Controlla sempre l'accuratezza.*

Valuta la precisione di questo riepilogo. 👍  👎

| Condividi riepilogo | **Modifica** |

Cordiali saluti,

Zoom
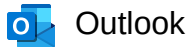
+1.888.799.9666

©2025 Zoom Communications, Inc.

Visita zoom.com

55 Almaden Blvd
San Jose, CA 95113

Outlook

**Meeting Summary for Course: Geocomputation and geospatial analysis**

**From** Meeting Summary with AI Companion <no-reply@zoom.us>
**Date** Thu 4/3/2025 12:43 PM
**To** Amatulli, Giuseppe <giuseppe.amatulli@yale.edu>

NEW  Meeting summary with AI Companion now supports additional languages in preview.

Learn More

## Meeting summary for Course: Geocomputation and geospatial analysis (04/03/2025)

### Quick recap

The meeting covered a wide range of topics related to bash commands, file management, and geospatial data processing using GDAL and other tools. Participants discussed various techniques for data manipulation, raster analysis, and the importance of understanding data types and no data values in digital image processing. The group also explored the use of virtual raster tiles for managing large datasets and concluded with a discussion on ongoing data processing tasks for a soil and water-related project.

### Next steps

• Fabio to download species presence data for the red swamp crayfish from iNaturalist and GBIF, expanding the search area to include the Mediterranean region, North Africa, and Turkey.
• Fabio to create a Jupyter notebook with scripts to download the species presence data.
• Fabio to download climatological, soil, and hydrological data as directed by Giuseppe.
• Fabio to perform data cleaning on the downloaded species presence data, removing points that fall in the sea or other inappropriate locations.
• Fabio to prepare a species distribution model for the red swamp crayfish.
• Fabio to arrive at the Lecce meeting with all presence and climate data downloaded, including at least 10 layers for testing.
• Saverio to review the SoilGrids data on the cluster and identify any missing tiles or variables.
• Saverio to rerun the SoilGrids download script if necessary to obtain missing data.
• Giuseppe to rerun the growing and cubeless scripts for the old SoilGrids data.

• Saverio to check if the water content variables (WWP1, WWP2, WWP3, and WWPS) are present in the SoilGrids data and redownload if needed.

---

## Summary

### Bash Commands and File Management

Giuseppe reviews basic bash commands and file management techniques with the group. He demonstrates how to use wget to download files from the internet, emphasizing its usefulness for static web pages but noting potential complications with authentication. Saverio adds that wget can be used in for loops to download multiple files efficiently. Giuseppe also covers changing directories, copying and moving files, and using the tmp folder for temporary storage. He introduces rsync as a more efficient alternative to copy for updating files. The group discusses the importance of understanding file paths and using commands like pwd and readlink to navigate the file system. Giuseppe mentions grep for searching within files and suggests finding a reference PDF for common bash commands.

### Command Line Tools for Data Manipulation

Giuseppe demonstrates how to use various command-line tools for data manipulation and searching, including grep, awk, and control+R for reverse searching command history. He explains that awk is useful for selecting lines or rows in large text files and manipulating data. Giuseppe also covers sorting files, joining files based on ID columns, and mentions that these commands often have built-in parallelization for efficiency. He encourages participants to be curious and search for commands to solve their specific needs. Chiara shares her experience with awk for simple data manipulations, and Fatemeh asks about separating characters in a text file, to which Giuseppe suggests using awk's gsub command or Python for more complex cases.

### Raster Data vs Vector Data

Giuseppe explains the advantages of raster data over vector data for certain computational tasks, noting that rasters are faster to process as they are organized in a matrix format. He then discusses several GDAL commands, focusing on gdalinfo, gdal_translate, and gdalwarp. Gdalinfo is used to display metadata of an image, gdal_translate can convert between formats and crop images, while gdalwarp is used for changing projections. Giuseppe emphasizes the importance of using EPSG codes for projections and recommends using GDAL for standardization across different software. Saverio adds that gdalwarp can also be used to change pixel resolution for upscaling or downscaling images.

### GDAL Commands in QGIS Demonstration

Giuseppe demonstrates how to use GDAL commands in QGIS, focusing on raster operations. He explains that many QGIS tools are based on GDAL functions, which can be accessed through the command line. Sofia asks about standalone Python scripting, and Giuseppe shows how to copy QGIS commands as Python code. The group discusses various GIS operations, including vector-to-raster conversion, zonal statistics, and DEM analysis, noting which can be done with GDAL and which require additional tools like GRASS or PKTools.

## Data Types and No Data Values

Giuseppe explains the importance of understanding data types and no data values in digital image processing. He discusses how different data types (byte, integer, float) affect file size and value ranges, emphasizing the need to choose the appropriate type for efficient storage and processing. Giuseppe also highlights the significance of correctly labeling and handling no data values, which are often represented by specific numbers at the extreme ends of data type ranges. He warns that mishandling no data values can lead to errors in analysis, statistics, and resampling. Giuseppe demonstrates the use of tools like GDAL and OpenEV for examining image properties and pixel values, and encourages careful consideration of data types and no data values when working with geospatial data.

## Raster Image Processing With GDAL

Giuseppe explains how to use GDAL commands to process raster images. He demonstrates using gdal_info to get general information about images, including minimum and maximum pixel values. He shows how to loop through multiple files to compute min/max values, which can help identify outliers. Giuseppe then covers gdal_calc for performing raster calculations, emphasizing the importance of ensuring input matrices have the same dimensions and coordinates. He explains how to use gdal_warp to change image resolution and crop files, noting the differences in coordinate ordering between commands. Giuseppe discusses resampling techniques like nearest neighbor and bilinear interpolation, advising when to use each method based on the data type (e.g. land cover vs elevation models).

## Raster and Vector File Analysis

In the meeting, Giuseppe discussed various aspects of raster and vector file analysis, including cropping, changing pixel resolution, and geolocation. He also introduced the concept of virtual raster tiles (VRT) and their usefulness in managing large datasets. Saverio added insights on the use of VRT for efficient data processing and visualization. Teodoro shared his experience using VRT for time series analysis. The team also discussed the importance of understanding pixel values and the potential of using the "ogr info" command for attribute table manipulation. Sofia asked about using tiling for moving window analysis, to which Giuseppe provided guidance on using the "pk filter" command. The team agreed to continue their discussions in the next meeting.

## Soil Data Processing and Backup Discussion

Giuseppe and Saverio discuss ongoing data processing tasks for a soil and water-related project. Giuseppe mentions he has restarted some scripts to regenerate missing tiles for water deficit and water carrying capacity data. Saverio notes additional soil parameters that might be useful, such as wilting point water content and bulk density. They agree to keep both the old and new versions of the soil grid data, with Giuseppe suggesting to rerun the scripts and download any missing files. Saverio confirms he has saved all necessary URLs and can easily download specific files if needed.

*AI-generated content may be inaccurate or misleading. Always check for accuracy.*

Please rate the accuracy of this summary.  👍  👎
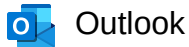
Share summary    Edit

Best,

Zoom

+1.888.799.9666

©2025 Zoom Communications, Inc.

Visit zoom.com

55 Almaden Blvd
San Jose, CA 95113

**Outlook**

---

**Meeting Summary for Course: Geocomputation and geospatial analysis**

---

**From** Meeting Summary with AI Companion <no-reply@zoom.us>
**Date** Tue 4/8/2025 12:02 PM
**To** Amatulli, Giuseppe <giuseppe.amatulli@yale.edu>

**NEW** Meeting summary with AI Companion now supports additional languages in preview.                                     **Learn More**

## Meeting summary for Course: Geocomputation and geospatial analysis (04/08/2025)

### Quick recap

Saverio led a Python session, addressing issues with accessing shared folders, opening files in Jupyter Lab, and using GDAL commands for GIS operations. He introduced Python and Jupyter notebooks, emphasizing their accessibility and extensive library support, and discussed the basics of NumPy arrays and their advantages over Python lists for numerical computations. The session also covered data cleaning and processing using numpy, the Python libraries NumPy and Pandas, and various Pandas operations for data manipulation and analysis.

### Next steps

• All attendees to import their own data sets (e.g. from Excel) into Python and practice using pandas for data manipulation.
• All attendees to try using artificial intelligence tools like ChatGPT, Perplexity, or GitHub Copilot for coding assistance and debugging.
• All attendees to prepare questions and topics for discussion at tomorrow's Q&A session.
• Sofia to follow up on creating a new shared folder in the virtual machine to resolve file access issues.
• Saverio to cover data visualization techniques using pandas in the next lesson.
• Giuseppe to point out information about Python environments in the syllabus.
• Chiara to practice using OpenEV and QGIS for visualizing raster data.
• Luca to explore using Climate Data Operator (CDO) for working with NetCDF files.

---

## Summary

### Python Session Troubleshooting and GDAL Commands

The summary discusses a Python session led by Saverio. Sofia encounters issues accessing shared folders in her virtual machine, which Giuseppe helps troubleshoot. Luca asks about opening files in Jupyter Lab, and Giuseppe explains the differences between using the bash console and Jupyter. Chiara has questions about using GDAL commands to analyze raster data. Giuseppe reviews GDAL and OGR commands for GIS operations and encourages participants to experiment and ask questions. He also mentions an upcoming catch-up session for more in-depth discussions. Saverio begins his portion by instructing participants on how to update their files and access the lesson materials.

### Python and Jupyter Notebook Overview

Saverio provides an overview of Python and Jupyter notebooks. He explains that Python is widely used due to its accessibility and extensive library support. The lesson aims to introduce the Jupyter notebook environment and basic libraries like pandas and NumPy for data analysis. Saverio discusses the benefits of Jupyter notebooks, including cell-by-cell execution and the ability to create interactive reports. He also mentions Google Colab as a cloud-based alternative for prototyping and experimenting with code. Saverio briefly touches on version control systems like GitHub and GitLab, emphasizing their importance in collaborative coding and tracking modifications. He then introduces NumPy as a powerful library for numerical computations and data manipulation, highlighting its widespread use in other Python libraries.

### NumPy Arrays for Efficient Computation

Saverio explains the basics of NumPy arrays and their advantages over Python lists for numerical computations. He demonstrates how to create arrays, perform operations on them, and access elements using indexing and slicing. He also covers multidimensional arrays, data types, and memory efficiency. Saverio shows examples of boolean operations, masking, and broadcasting with NumPy arrays. He emphasizes the importance of choosing appropriate data types to optimize memory usage and processing speed, especially when working with limited hardware resources.

### Data Cleaning and Processing Techniques

Saverio discussed data cleaning and processing using numpy, emphasizing the importance of handling missing data values. He demonstrated techniques for identifying and replacing missing data, as well as removing outliers. Saverio also highlighted the use of numpy's documentation for understanding module structures and parameters. He concluded by suggesting that for simple computations, Rwk or bash might be faster, but for more complex operations, Python could be more suitable.

### NumPy and Pandas Library Overview

Saverio introduced the Python libraries NumPy and Pandas, explaining their functionalities and how they work. He demonstrated how to create, modify, and import structured data using these libraries. He also discussed the importance of understanding the logic behind these libraries, as they have a lot of mechanics running in the background. Saverio also touched on the concept of labeled data management in Pandas, which is crucial for data handling. He concluded the session by explaining how to create a data frame from a dictionary.

**Pandas Operations for Data Manipulation**

The segment discusses various Pandas operations for data manipulation and analysis. Saverio demonstrates how to create, clean, and explore data frames, including importing CSV files, handling missing data, and performing basic statistical operations. He also shows how to use pivot tables for data summarization and how to export modified data to CSV files. The discussion emphasizes the importance of data cleaning and standardization for effective analysis and automation of data processing workflows.

**Pandas for Excel Data Import**

Giuseppe and Saverio discuss importing data from Excel into Python using Pandas. They mention that Pandas can read Excel files directly, but sometimes conversion to a text format like CSV may be necessary. Sofia notes there is a read_excel function available. Saverio shares his experience working with heterogeneous datasets for a landslide project, emphasizing the importance of data cleaning and standardization. Giuseppe encourages everyone to practice importing and working with datasets before the next meeting. They also discuss using AI tools like ChatGPT and GitHub Copilot for coding assistance and debugging.

*AI-generated content may be inaccurate or misleading. Always check for accuracy.*

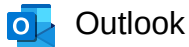Please rate the accuracy of this summary.   👍   👎

| Share summary | Edit |

Best,

Zoom

**Outlook**

---

**Riepilogo riunione per Course: Geocomputation and geospatial analysis**

---

**From** Meeting Summary with AI Companion <no-reply@zoom.us>
**Date** Wed 4/9/2025 10:20 AM
**To** Amatulli, Giuseppe <giuseppe.amatulli@yale.edu>

**NEW** Riepilogo della riunione con AI Companion ora supporta ulteriori lingue in anteprima.    **Ulteriori informazioni**

## Riepilogo della riunione per Course: Geocomputation and geospatial analysis (04/09/2025)

### Riepilogo rapido

La riunione ha esplorato l'importanza dell'utilizzo di Python per l'analisi geospaziale e la visualizzazione dei dati, incoraggiando gli studenti a sperimentare e chiedere aiuto durante le sessioni pratiche. Sono stati discussi vari aspetti tecnici dell'utilizzo di strumenti GIS, tra cui l'importazione di dati, la risoluzione di problemi e l'estrazione di statistiche. Infine, sono stati forniti consigli sull'uso di software come QGIS e Python per i progetti degli studenti, con una panoramica della prossima lezione che includerà esempi pratici di analisi dati e elaborazione di immagini satellitari.

### Fasi successive

• Sofia: Riproiettare lo shapefile sulle coordinate del raster e utilizzare GDAL extracto per estrarre le statistiche dalle zone dello shapefile.
• Luca: Estrarre alcune bande specifiche dai file NetCDF e salvarle in formato TIF.
• Giusy: Seguire i video tutorial su GRASS GIS per l'analisi idrologica e la creazione di sottobacini.
• Alessandro: Provare a trasferire il lavoro fatto in R utilizzando le librerie Python presentate per l'analisi geospaziale.
• Tutti i partecipanti: Rivedere le registrazioni delle lezioni precedenti sulla pagina YouTube del corso.
• Tutti i partecipanti: Prepararsi per la prossima lezione pratica con Saverio sugli esempi di analisi dati e geospaziale in Python.
• Rocco_Fracchiolla: Registrarsi alla piattaforma Itineris utilizzando l'email fornita al corso.
• Giuseppe: Aggiornare la pagina web del corso con il link corretto alla registrazione della lezione 3.

**Riepilogo**

**Esercitarsi Con Codici Python.**
Giuseppe, Alessandro e Saverio discutono l'importanza di esplorare e sperimentare con il codice Python per l'analisi geospaziale e la visualizzazione dei dati. Saverio spiega che gli studenti possono utilizzare gli script forniti come punto di partenza, modificandoli per adattarli alle proprie esigenze, e sottolinea i vantaggi di lavorare con librerie come Rasterio, Geopandas e Matplotlib per creare grafici e mappe avanzate. Il gruppo incoraggia gli studenti a essere curiosi, provare il codice e chiedere aiuto durante le sessioni pratiche per superare le limitazioni dei software tradizionali come Excel e QGIS.

**Risolvere Problemi Di Riproiezione**
Sofia discute l'importazione di una cartella condivisa e il tentativo di sovrapporre un raster a un file shape per estrarre statistiche. Giuseppe e Saverio offrono consigli su come risolvere i problemi di proiezione e formato dei file, suggerendo l'uso del comando "gdalinfo" per ottenere informazioni sui file e verificare le coordinate di riferimento. Il gruppo analizza gli errori riscontrati durante la riproiezione dei file e discute possibili soluzioni.

**Utilizzo Di Strumenti GIS.**
Giuseppe guida Sofia e Saverio attraverso un processo di configurazione e utilizzo di strumenti GIS, correggendo errori comuni come l'inversione di input e output. Discutono l'uso di formati file come GeoPackage invece di Shapefile, e Giuseppe consiglia di utilizzare sempre l'opzione "-overwrite" quando si lavora con OGR. Il gruppo esplora vari comandi e opzioni per la manipolazione dei dati geografici, con Giuseppe che condivide consigli basati sulla sua esperienza ventennale nel campo.

**Uso Di Strumenti GIS**
Durante la riunione, Giuseppe e Sofia discutono l'uso di strumenti GIS come GDAL e Picket Tours per l'estrazione di statistiche da immagini raster utilizzando poligoni vettoriali. Saverio menziona una prossima lezione con esempi pratici su questo argomento, mentre Luca pone domande sull'utilizzo di GDAL Info e l'estrazione di bande specifiche da file raster. La discussione si conclude con un problema tecnico riscontrato da Luca nell'utilizzo combinato di dati raster e vettoriali in un unico file NetCDF.

**Analisi Dati Spaziali Con Python.**
Giuseppe illustra l'uso di software GIS e analisi di dati spaziali, consigliando agli studenti di utilizzare QGIS e Python per i loro progetti. Saverio presenta una panoramica della prossima lezione, che includerà esempi pratici di analisi dati con Python, elaborazione di immagini satellitari e operazioni su dati vettoriali. Gli studenti pongono domande su registrazioni delle lezioni, piattaforme online e l'uso di diversi linguaggi di programmazione per l'analisi dei dati.

*Il contenuto generato da intelligenza artificiale potrebbe essere impreciso o fuorviante. Controlla sempre l'accuratezza.*

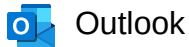Valuta la precisione di questo riepilogo. 👍 👎

| Condividi riepilogo | **Modifica** |
|:---:|:---:|

Cordiali saluti,

Zoom

+1.888.799.9666

©2025 Zoom Communications, Inc.

Visita zoom.com

55 Almaden Blvd
San Jose, CA 95113

![o] **Outlook**

---

**Meeting Summary for Course: Geocomputation and geospatial analysis**

---

**From** Meeting Summary with AI Companion <no-reply@zoom.us>
**Date** Thu 4/10/2025 11:51 AM
**To** Amatulli, Giuseppe <giuseppe.amatulli@yale.edu>

**NEW**   Meeting summary with AI Companion now supports additional languages in   **Learn More**
          preview.

## Meeting summary for Course: Geocomputation and geospatial analysis (04/10/2025)

### Quick recap

Saverio introduced the use of matplotlib and seaborn libraries for data visualization in Python, demonstrating various plots and emphasizing the importance of histograms for understanding data distribution. He also discussed the use of Rasterio for handling raster data, including its integration with other libraries, and the importance of correctly labeling data in software. The team also discussed the use of shapefiles and geopackages in data processing, the utility of certain tools for data analysis, and the potential of these tools for creating buffers, cropping, and populating data sets from raster or other shapefiles.

### Next steps

- All attendees to arrive in Lecce with a plan and dataset ready for the in-person meeting.
- All attendees to check the information sent by Mario regarding the class location and logistics.
- Mario to provide information about bus schedules for transportation to the campus area.
- Attendees with cars to coordinate carpooling to the campus.
- Giuseppe and Saverio to arrive early Tuesday morning for the in-person meeting.
- All attendees to participate in social dinners and brainstorming sessions outside of class time.
- All attendees to prepare questions or ideas for applying the Python techniques discussed (e.g. transect analysis, raster processing, shapefile manipulation) to their own datasets.

---

## Summary

### Matplotlib and Seaborn for Data Visualization

Saverio introduces matplotlib and seaborn libraries for data visualization in Python. He demonstrates how to create various plots including line plots, histograms, scatter plots, and box plots using data from a Matera weather dataset. Giuseppe emphasizes the importance of histograms for understanding data distribution and identifying outliers during initial data exploration. Saverio also shows how to use seaborn for more advanced visualizations like kernel density plots and heatmaps. He highlights the extensive documentation and online resources available for these libraries, encouraging participants to explore and adapt example code for their own datasets. The session concludes with a brief introduction to geospatial data handling using the Rasterio library.

### Rasterio for Raster Data Manipulation

Saverio discussed the use of Rasterio for handling raster data, including its integration with other libraries like NumPy and SciPy. He explained how to import necessary libraries, access files, and plot raster data. Saverio also demonstrated how to perform basic manipulations like cropping and transforming raster data, and how to handle metadata. He showed how to use Rasterio for filtering and masking raster data, and how to perform operations like NDVI calculation. Saverio emphasized the importance of correctly handling no data values and updating metadata. He also mentioned the possibility of using Rasterio for more complex operations like object detection.

### Data Labeling and Python Libraries

Giuseppe discussed the importance of correctly labeling data in software, emphasizing that this would likely prevent issues in older software and new data. He also mentioned the flexibility of Python libraries, such as numpy, for complex filtering operations. Saverio added that numpy's power lies in its ability to access and manipulate individual information in each array, but this can be memory-intensive. He suggested that downsizing could be more efficiently done on Gidal. The team also briefly discussed GeoPandas, a library used for handling vector files, but no specific decisions or next steps were mentioned.

### Shapefiles and Geopackages in Data Processing

Saverio discussed the use of shapefiles and geopackages in data processing. He explained how to read and manipulate shapefiles, including extracting specific information and performing operations like intersection and cropping. Saverio also demonstrated how to create a new shapefile from a raster file, using the bounding box of the raster to select relevant polygons from the original shapefile. He further showed how to calculate the average NDVI value for each selected municipality and plot the results. Saverio also discussed the use of shapely geometry to create a polyline from a list of points and how to extract NDVI values along this polyline. He concluded by demonstrating how to create a buffer area from a polyline and use it to crop a raster file. Giuseppe suggested that attendees could use these techniques for their own image and vector file processing tasks.

### Raster Data Analysis Tools Discussed

In the meeting, Saverio and Giuseppe discussed the utility of certain tools for data analysis, particularly in relation to raster data. They highlighted the potential of these tools

for creating buffers, cropping, and populating data sets from raster or other shapefiles. Giuseppe emphasized the versatility of these tools, especially in creating reports and presentations. The team also discussed the possibility of applying these tools to time series data and the importance of planning ahead for the upcoming in-person meeting.

*AI-generated content may be inaccurate or misleading. Always check for accuracy.*

Please rate the accuracy of this summary.  👍  👎

[ Share summary ]   [ Edit ]

Best,

Zoom

+1.888.799.9666

©2025 Zoom Communications, Inc.

Visit zoom.com

55 Almaden Blvd
San Jose, CA 95113