

## Project 2

The **Birds Occurrence Data in Kenya** dataset captures bird sightings across the Republic of Kenya, with the purpose of understanding the **national abundance** of various bird species. The dataset spans from **June 1, 2017, to May 16, 2018**, covering both the **wet** and **dry seasons**, as bird counting is conducted twice a year. Bird observations are collected using **transects** within specific focal areas, ensuring a systematic approach to data collection. The data is curated and processed by experts from the **National Museums of Kenya** to ensure quality, with proper training provided to participants on bird counting techniques. This dataset provides valuable insights into bird distributions across Kenya.

Here is the link to the page that provides a detailed description of the dataset and includes the download link: <https://www.gbif.org/dataset/d1e8a7af-097a-4ca8-aa91-62d3ff834711>

Use this dataset to answer **at least two** of the following questions. Addressing more than two may earn additional points, but this will depend on the **quality** and **depth** of your analysis.

In addition to submitting a PDF of your report, you must also include any Python code you developed for your analysis, as all analysis must be conducted using Python. Please submit your project through Moodle no later than **12 November 2024** and ensure both the report and code are included in your submission (put the code in a single file).

- (1) Analyze how bird species belonging to different taxonomic **orders** are distributed across Kenya. Some species may exhibit **clustering**, where they are more frequently found near each other in specific regions, while others may be more **dispersed**, spread more evenly across the country. Compare the results for different orders and determine whether similar patterns of clustering or dispersion are observed across all orders. Finally, explain if certain orders show unique distribution patterns.
- (2) Identify which regions in Kenya exhibit significant **clusters of bird sightings**, regardless of bird species. Use a map to visually display these clusters, highlighting areas with higher concentrations of sightings. Investigate whether there are any areas where bird sightings significantly differ from surrounding regions.
- (3) How do **elevation, precipitation, land cover**, and **distance from major water bodies** (like seas, lakes, and rivers) (choose two of these four) influence the overall **abundance of bird species** across Kenya? Conduct an analysis to assess how these environmental factors impact bird populations and explain the relationship between the variables you choose and bird abundance. Use maps or graphs to illustrate your findings.
- (4) A species richness map is a spatial representation that shows the number of different species (i.e., species richness) present in specific geographic areas by counting the unique species observed within that area. It highlights the distribution of biodiversity across a landscape by depicting where higher or lower concentrations of species occur.

**Course Title:** Environmental Data Analytics.

**Degree Program:** Master's in data science.

**Instructor:** Mohammad Mahdi Rajabi



UNIVERSITÉ DU  
LUXEMBOURG

---

Your task is to create a bird species richness map for Kenya. Your map should provide a clear visual representation of the geographic distribution of bird diversity and offer insights into potential biodiversity hotspots or areas of lower species richness. After doing so, analyze how bird species richness varies across the country.