INDU: Spread of the purple emperor in Sweden

Lars Arvestad

Contents

Introduction

The purple emperor butterfly (sälgskimmerfjäril) can be found in many parts of Europe, but has until recently only been seen occasionally in the southernmost parts of Sweden. In the past several years, however, it has begun to spread north. In this project, you will investigate its spread over Sweden in the past 25 years with the help of observation data from Artdatabanken (a site from SLU, Sveriges lantbruksuniversitet).

If you choose to do this project, you can receive a grade in the range E-C. For general requirements for projects and grade criteria, see the general instructions in *INDU—Individual Assignment in Programming*.

Task (E-C)

You will write a Python program that reads observation data for the purple emperor butterfly (see section "Data" below) and analyzes how its spread has changed over time. The program should create figures with the help of `matplotlib`.

The questions we want to answer are:

- **At what rate has the purple emperor spread northward?** Show in a figure how the northernmost observation has changed over time. That is, create a figure with years on the x-axis and northernmost latitude on the y-axis. Also, calculate the average latitude of observations for each year and show in a figure how it has changed over time.
- **How has the number of observations changed over time?** Show in a figure how many observations were made each year in Sweden.
- **When did the purple emperor arrive in Gotland?** Have your program search for observations made in Gotland in order to print out a *table* (using `print`) with columns "Year" and "Number of observations" collecting the number of observations per year.
- **What time of the year is the purple emperor active?** Use the observations to determine what time of year one can find a purple emperor in nature, and investigate whether this depends on the latitude. Your program should make present its determination in text. You are welcome to—but are not required to—present the data as a figure instead if you prefer.

Requirements

- In your figures, the axes should be labeled (for example “Year” and “Number of observations”)
- Your program should answer all the questions clearly and legibly.
- Your program should save the figures in PDF format with descriptive file names.

Report

As described in the general instructions, the report should focus on the program’s design and structure and should be at most three A4 pages. For this program, however, you can add an appendix containing the figures, answers to the questions, and explanations if necessary.

Data

In the project folder on the course website, you can find the file `25_years_of_salgskimmer.csv`, which contains four comma-separated columns with information on observations of the purple emperor. The two first columns are latitude and longitude, which describe where the observation took place (though hardly with the implied precision). The third column contains information on the date and time in a somewhat mysterious format. Take for example `2021-07-28T22:00:00.0000000Z`: the date is in the standard format, but separated from the time with the character `T`. The time is also in the standard format, but with a bizarre amount of precision and terminated with the letter `Z`. The fourth column contains the number of butterflies reported in the observation.

In total, the data contains 5239 observations. This is a filtered version of the data you can download from SLU’s site `https://www.artdatabanken.se`. Irrelevant data and personal information (the observers’ names) have been removed.

Grading

Details can be found with the general instructions in *INDU—Individual Assignment in Programming*.