

A close-up photograph of a butterfly with orange and black wings resting on a green leaf. The background is dark and out of focus, showing more foliage.

The butterfly effect

A Case for Biodiversity using Machine Learning

What are we looking at?



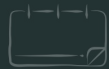
6,392,186
butterflies sightings



2001 - 2020



United Kingdom



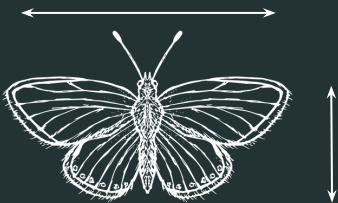
5,832,272
butterflies sightings

2004 - 2020

United Kingdom

Dataset 1 is merged with:

WIP:



Dataset 2

Butterflies Traits
(wingspan, flight duration...)



Dataset 3

List of endangered
butterfly species in the UK



Dataset 4

External factors
(temperatures, air quality...)

What do we want to predict?

Using a **Time Series model**,

an **estimation of the butterfly population evolution** over the next years

In the context of **climate change**

Why should we care?

The World Economic Forum states that **Biodiversity** is 'critically important' for 5 reasons, as it:

1. Ensures health and food security
2. Helps fight disease
3. Benefits business
4. Provides livelihood
5. Protects us

Why butterflies? The short life cycles are thought to be one of the best indicators of how healthy an environment is.

Initial findings

Dataset preprocessing issues

D1 (Sightings)



D2 (Traits)



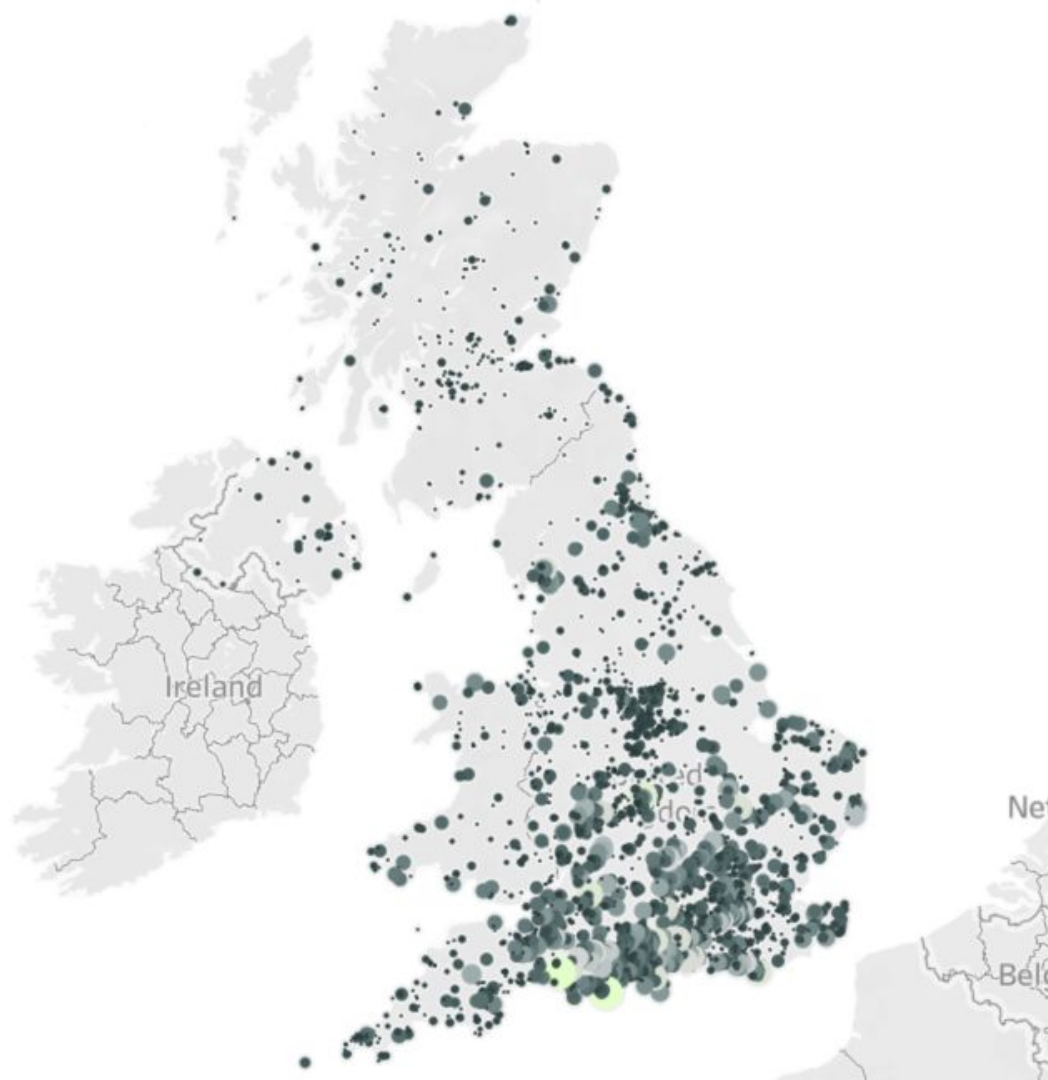
Merging D1 and D2

- Key = Species name
- Duplicates in D2
- Missing values in D2

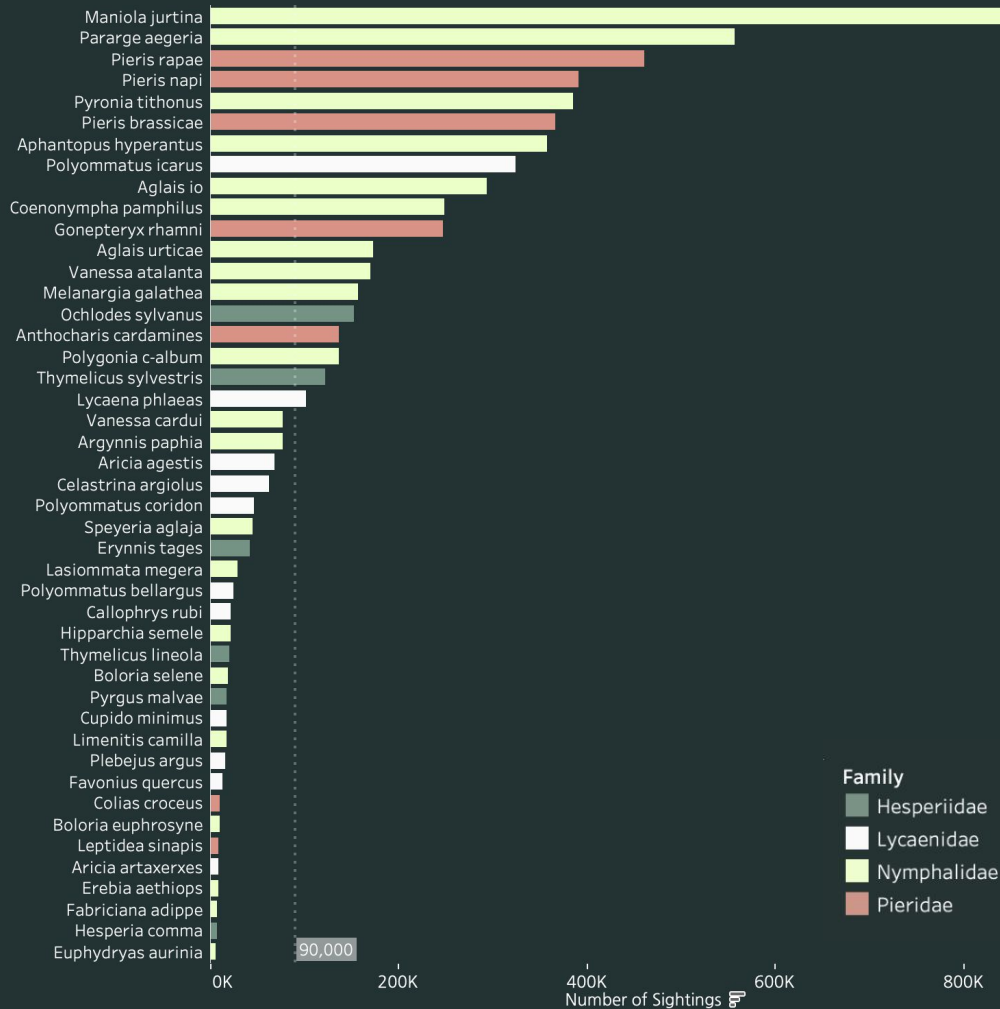
Sightings VS Surveys

- Missing location issues in first dataset
- Finding the number of surveys

UK sightings

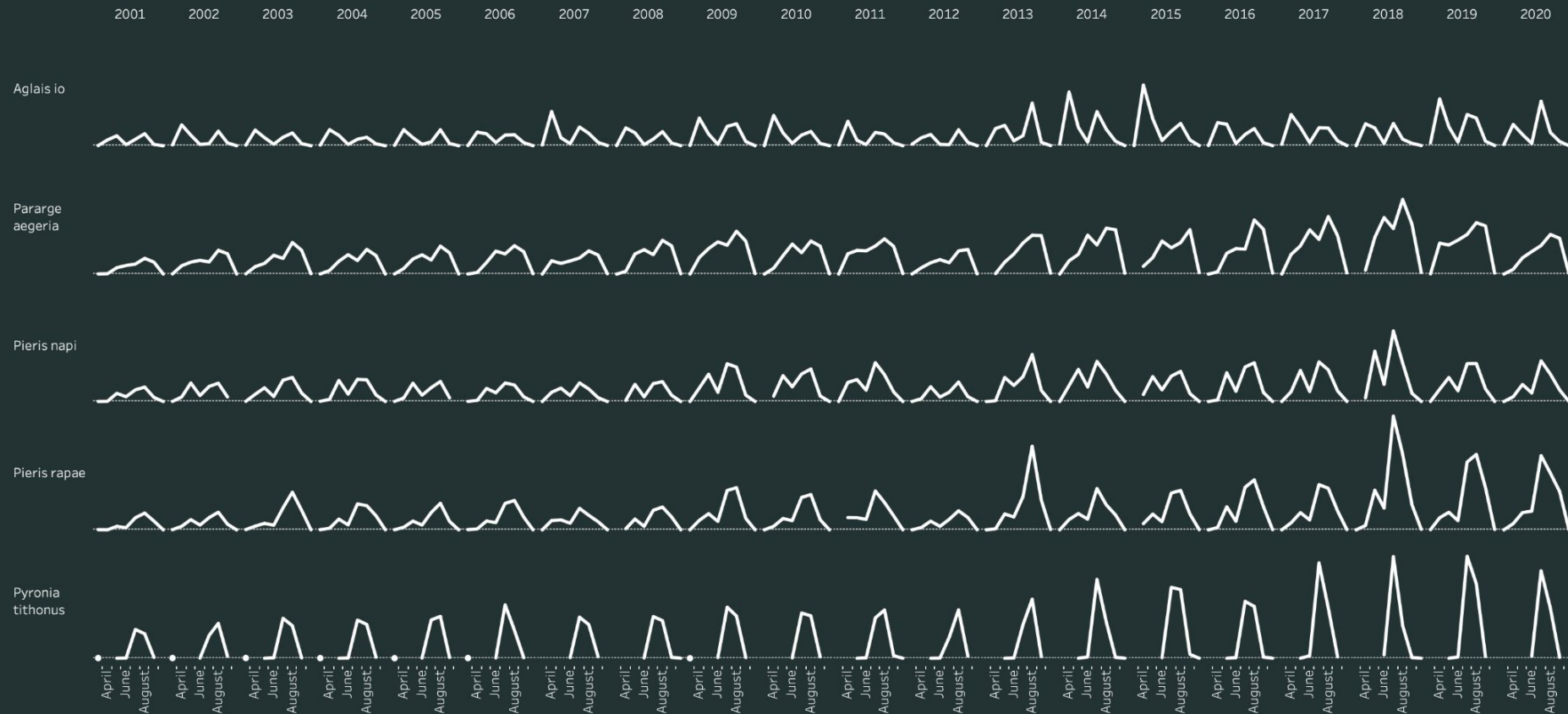


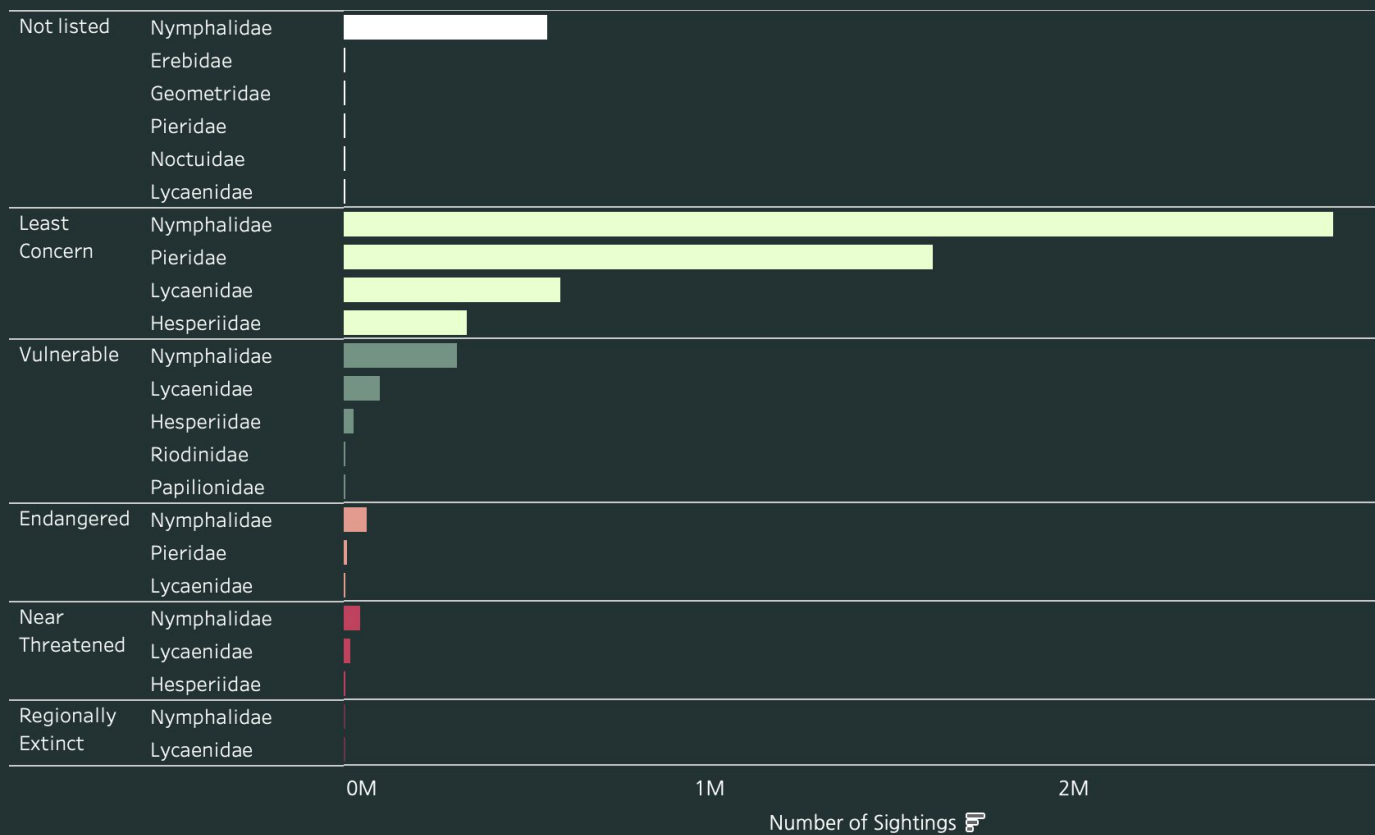
Most of the butterflies
sightings in England



We find 71 different butterfly species in the dataset, with an average of 90K sightings per species.

Butterflies sightings and their seasonal pattern for the 5 most seen species

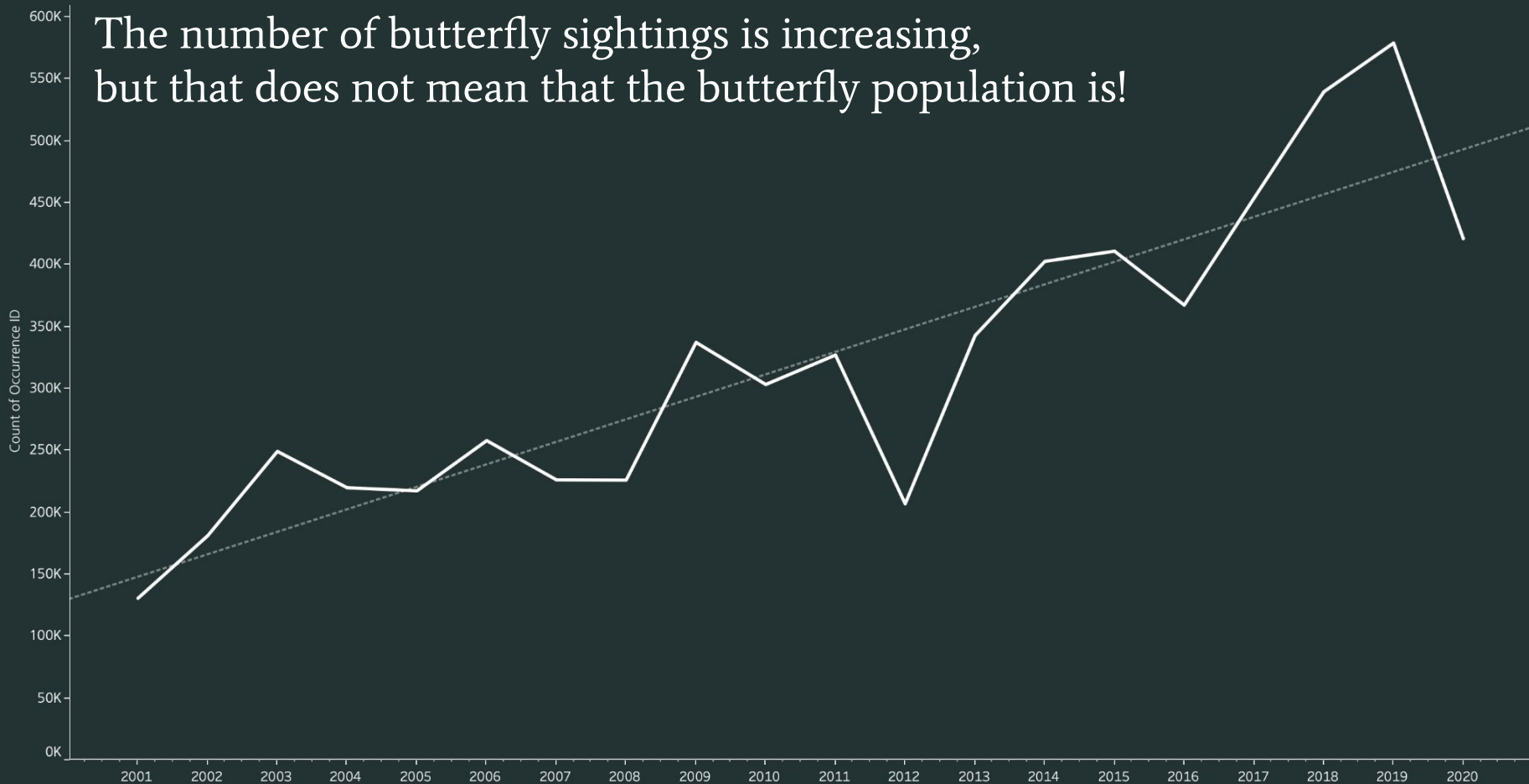


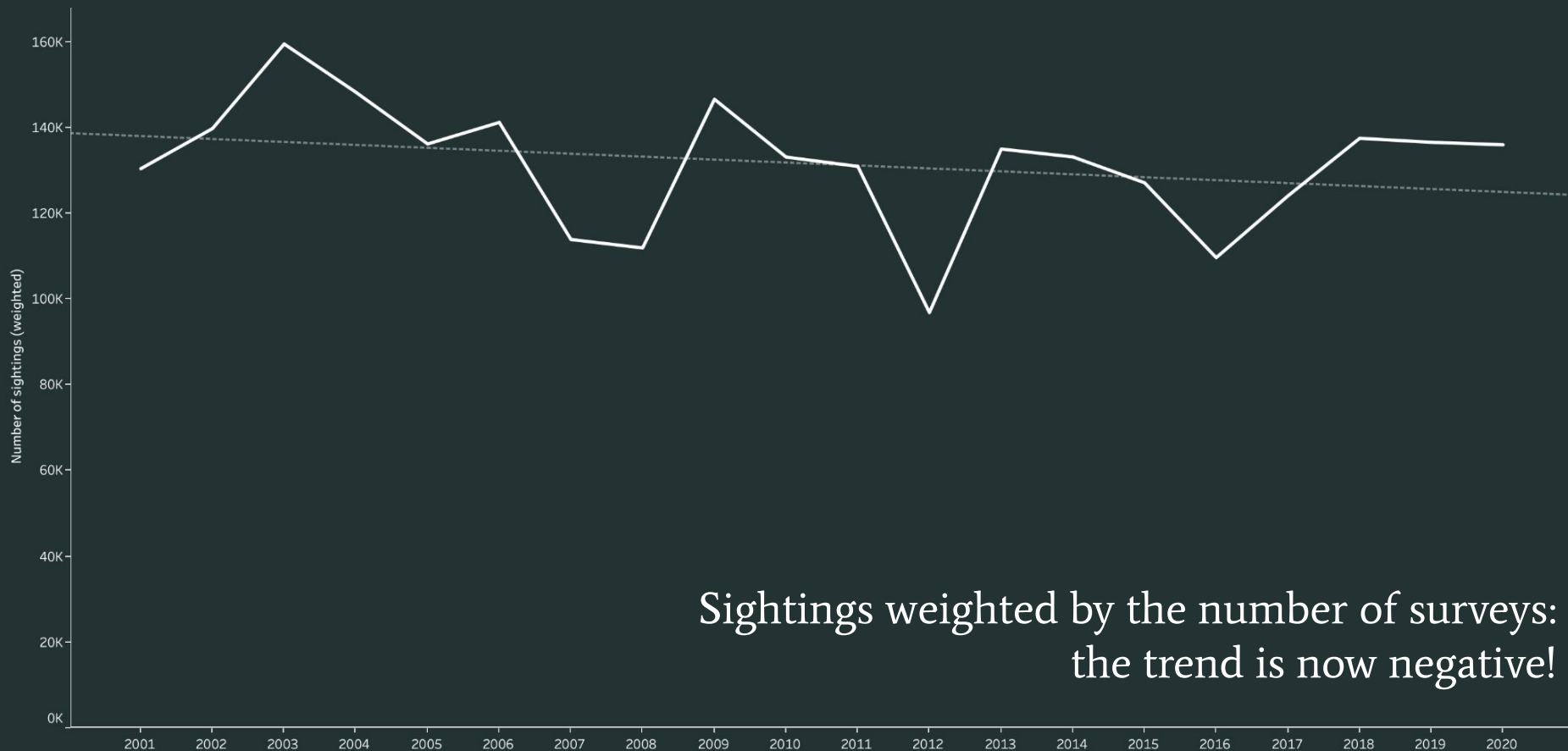


Sightings for the most vulnerable species

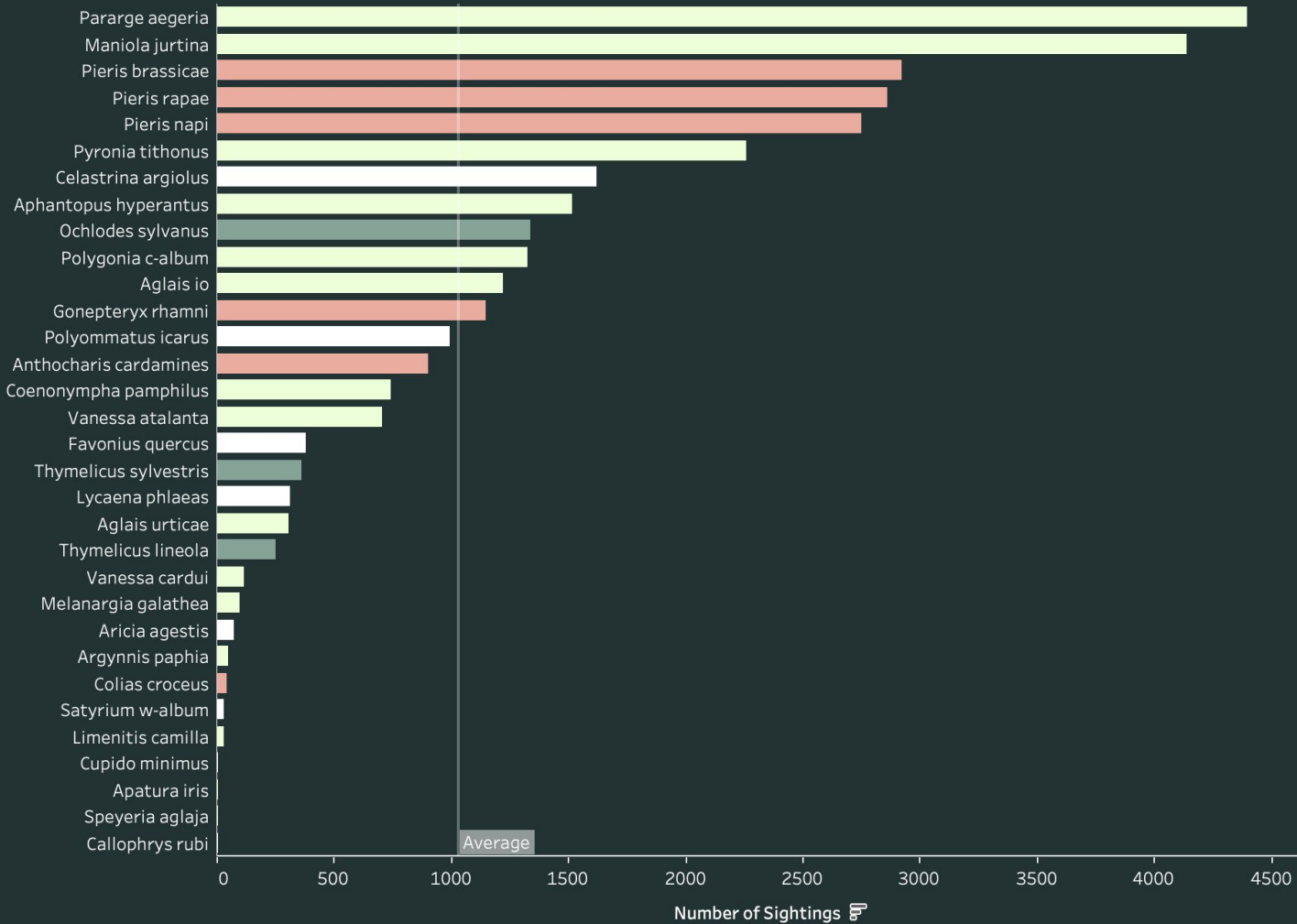
Surveys VS Sightings

The number of butterfly sightings is increasing,
but that does not mean that the butterfly population is!



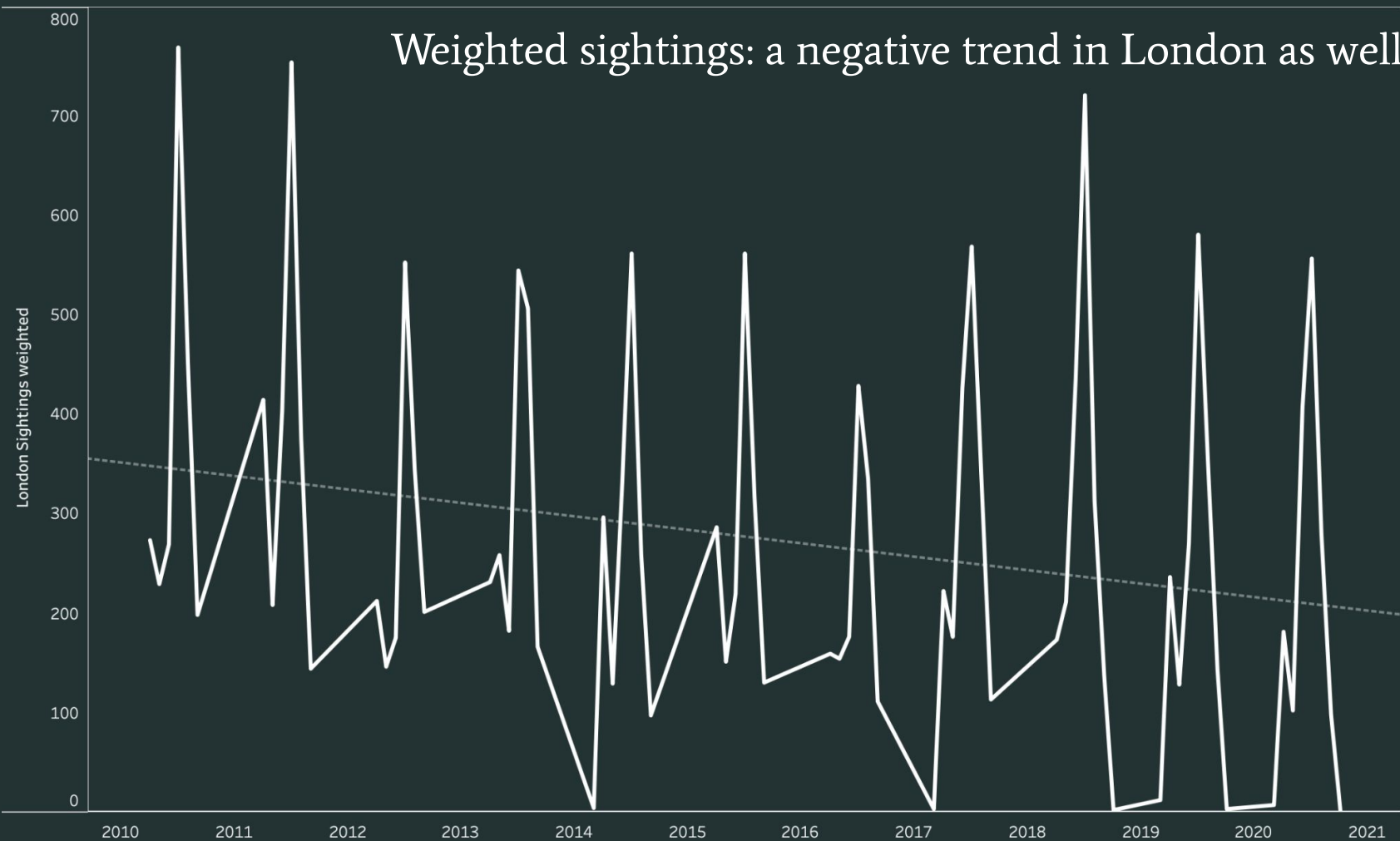


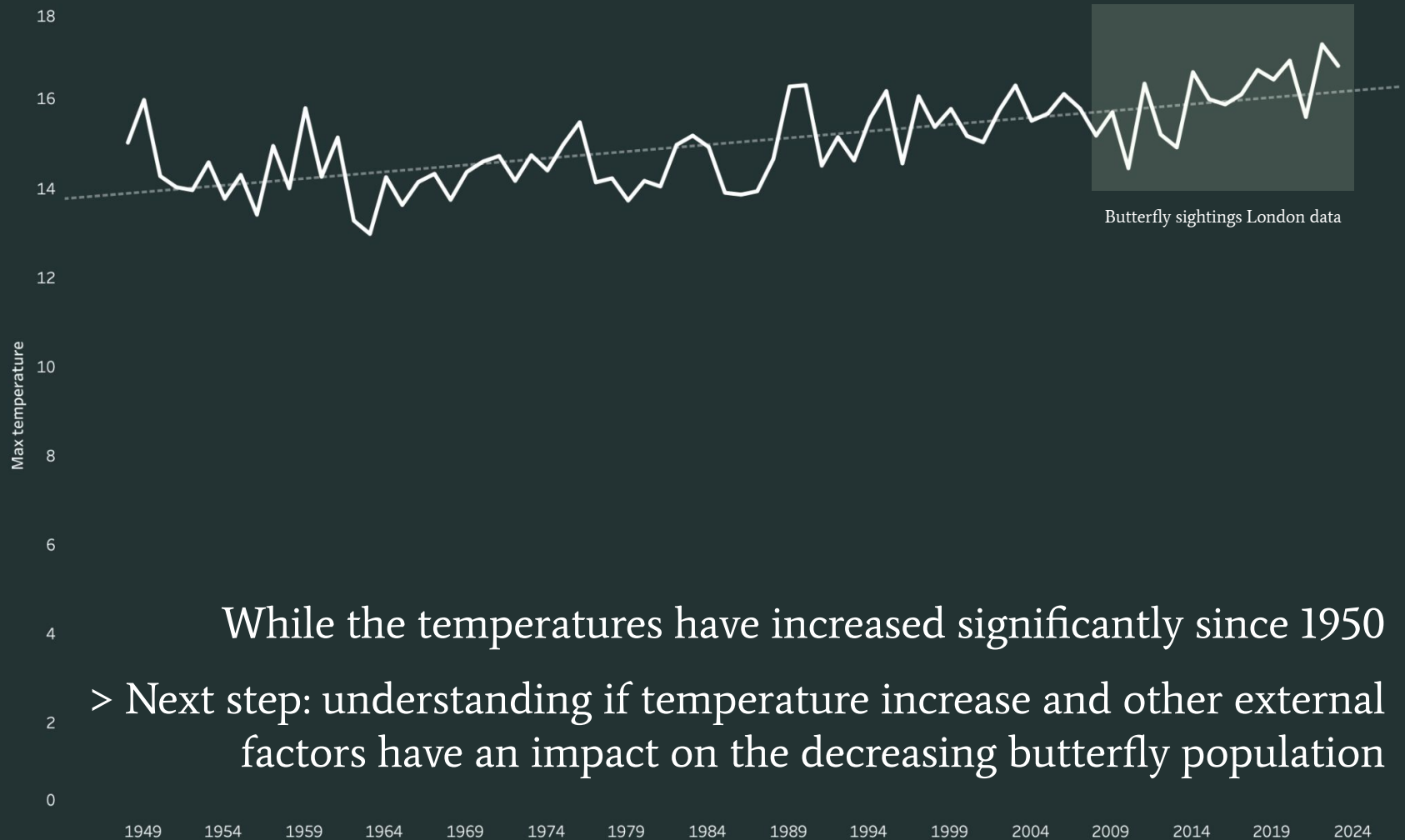
London sightings



We find 32 different butterfly species in London, with an average of 1K sightings per species.

Weighted sightings: a negative trend in London as well





Next steps

Next steps

- > Focus on sightings in one location – London
- > Time series model: butterfly population evolution predictions
- > Understand if and how external factors or Butterfly Traits
 impact the evolution historic data and future predictions

Thanks!