

Galaxy-Bricks a Tool for Data Literacy and Scientific Approach Education in the Context of Citizen Science



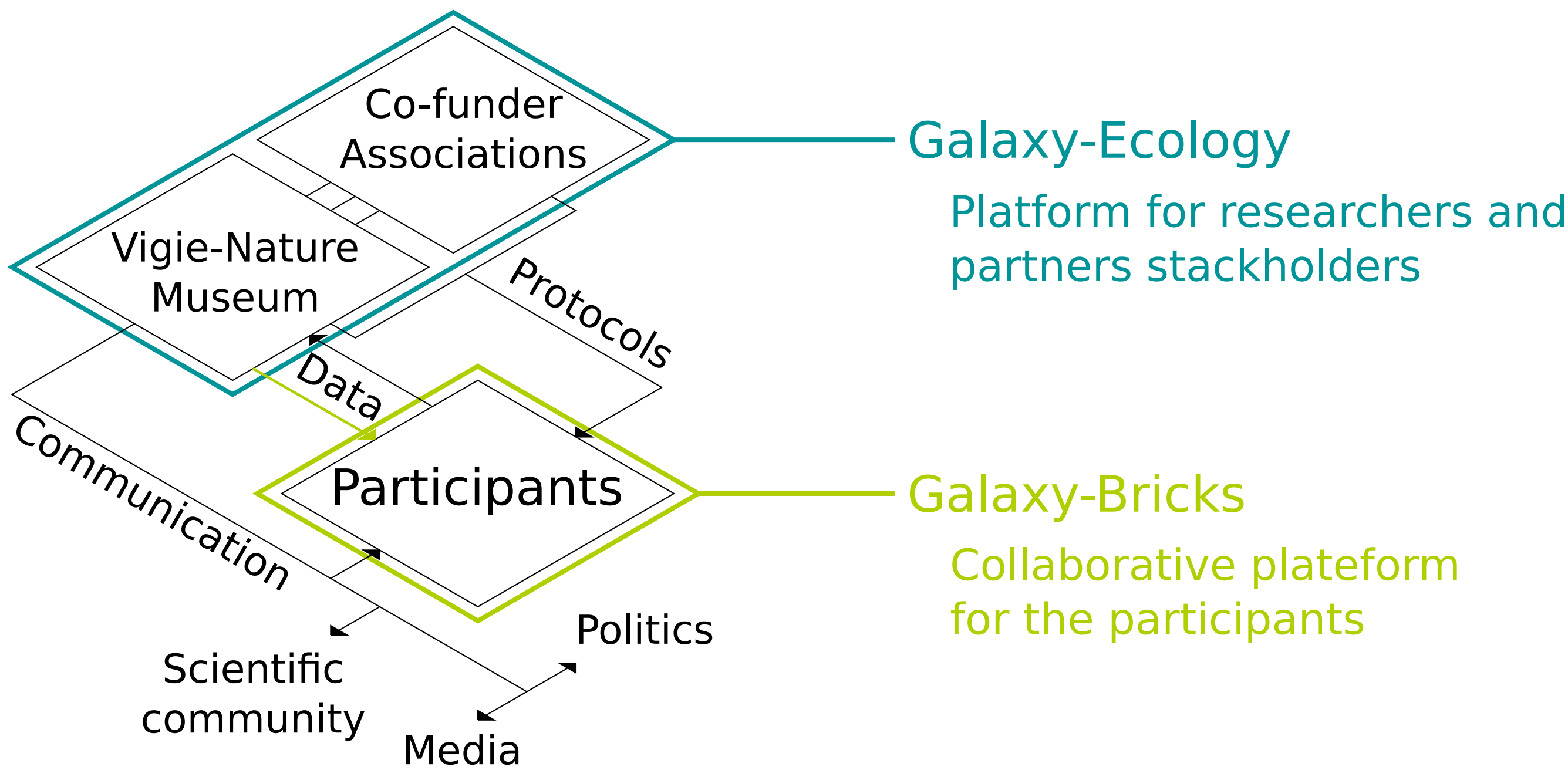
Simon Bénateau¹, Sébastien Turpin¹, Yvan Le Bras²

1. MNHN, CESCO, UMR 7204, 43 rue Buffon, 75005 Paris, France
2. MNHN, PatriNat, UMS 2006, station marine de Concarneau, 29900 Concarneau



Introduction

To increase the accessibility to data and extend collaborations between research teams and participants of our citizen science programs, we are developing Galaxy-Bricks, a data analysis tool based on Galaxy for Ecology. This platform will include user-friendly tools together with specific training material, designed for non-professional contributors. We are using the possibilities that Galaxy provides to reach this goal and we are investigating simplification of the graphical user interface using notably external softwares (Scratch and NGPhylogeny.fr).



Step 1 : Your research question
Effect of climate on insects

Step 2 : Import data

Step 3 : Process data

Group lines and summarise data

Parameters

1- Choose dataset
Pollinator data

2 - Choose which columns will be used to group the lines
Choose your column(s)
Observation number
Mean annual temperature
5 pears
Abundance

3- Selectionner l'opération à réaliser
Choose the column to summarise
Abundance

Choose one or several operation to run
Sum
Mean
Median
Count
Standard-deviation
Minimum
Maximum
Quartiles

Help 'Group lines and summarise data'

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum

Run

Your data

Q Pollinator data

Q Pollinator and climate data

Localisation	Mean annual temperature	Mean annual precipitations
Paris5	10	700
Paris6	10	700
Paris7	10	700
Paris8	10	700
Paris9	10	700
Paris10	10	700

Understand the data

Location

Mean annual temperature

Mean annual precipitations

Tools

Import data

Summarise data

Create or modify columns

Calculate

Visualise

Statistical tests

Linear regression with x

Compare means for

Calculate confidence interval

Khi2 test for

Import vigie nature data

Import climate data

Import climate data

Group lines on

use mean and

Scatterplot with x

Linear regression with x

Mean annual temperature join on Location

Mean annual precipitations join on Location

Mean annual temperature and Mean annual temperature

Standard deviation of Abundance to summarise data

Mean annual precipitations and y Mean abundance

Mean annual precipitations

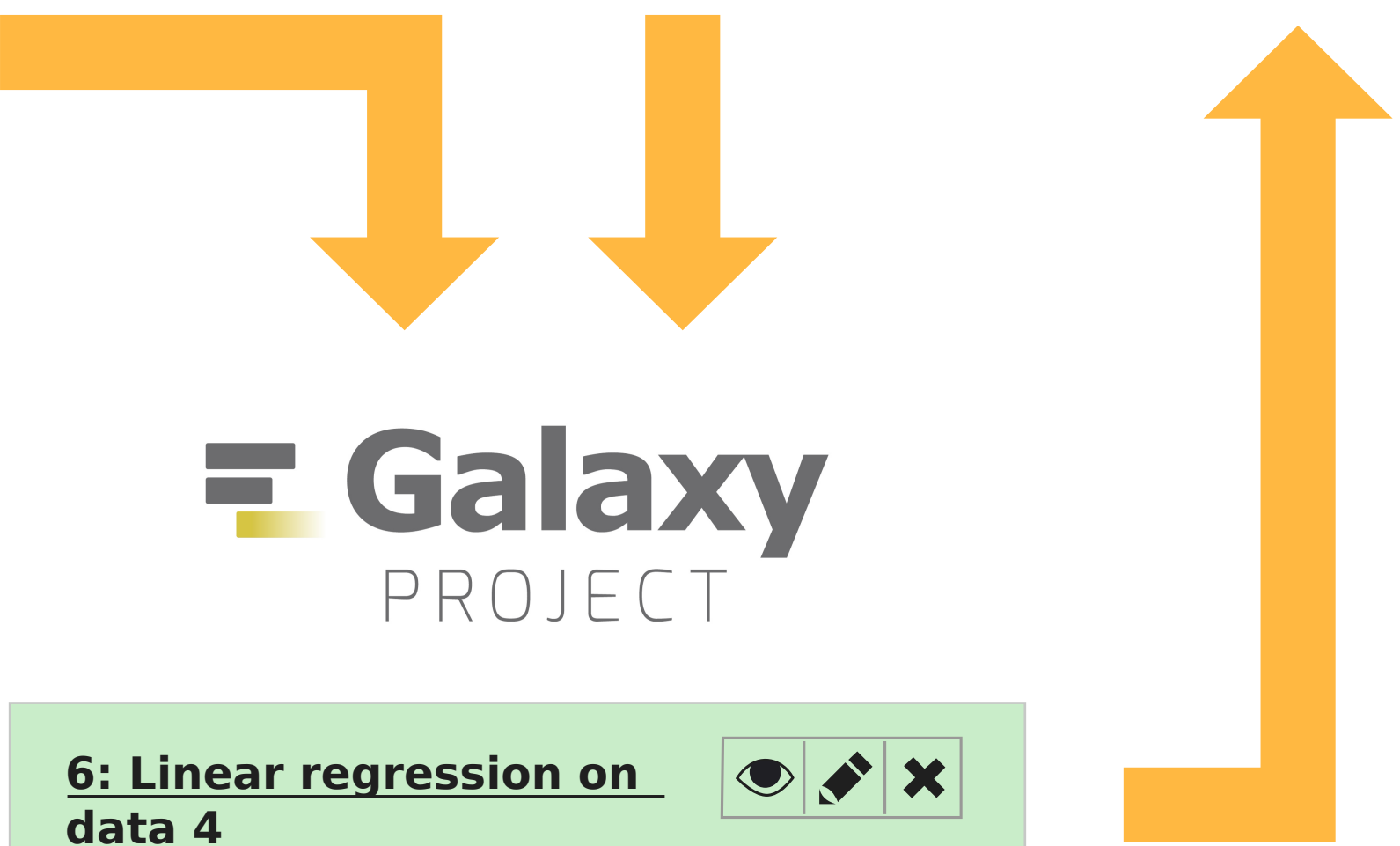
Results of linear regression on group and summarise data on pollinator and climate data

Equation:
Mean abundance = -0.01 x Temperature + 10.48
R2 = -0.02 Weak relation
Non significant result

Step 4 : Visualise data

Step 5 : Statistical tests

Step 6 : Conclude and edit report



Based on Galaxy

- to move towards FAIR practices
- to give access to the data the citizen produce
- to give access to high performance computing
- to reduce coding effort and share the tools produced with the community
- to allow the production of different user interfaces

6: Linear regression on data 4

5: Scatterplot w ggplot2 on data 4

4: Datamash on data 3

3: Join on data 1 and data 2

2: Get climatic data

1: Get Vigie-Nature data

Develop user interfaces

- to fit the diversity of participants (experts, amateurs, highschool students) by reducing the complexity of galaxy
- to give access to a user friendly interface facilitating scientific approach education