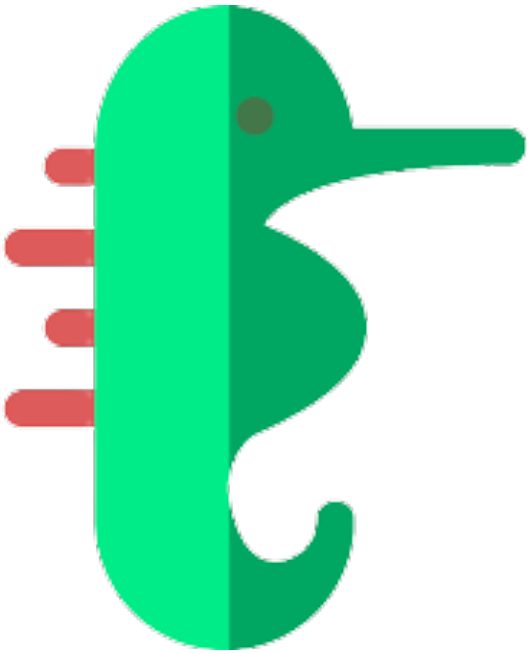


SeadragonSearch Analytics Documentation

None

Dominic Cain

None

Table of contents

1. Welcome to SeadragonSearch Analytics Documentation

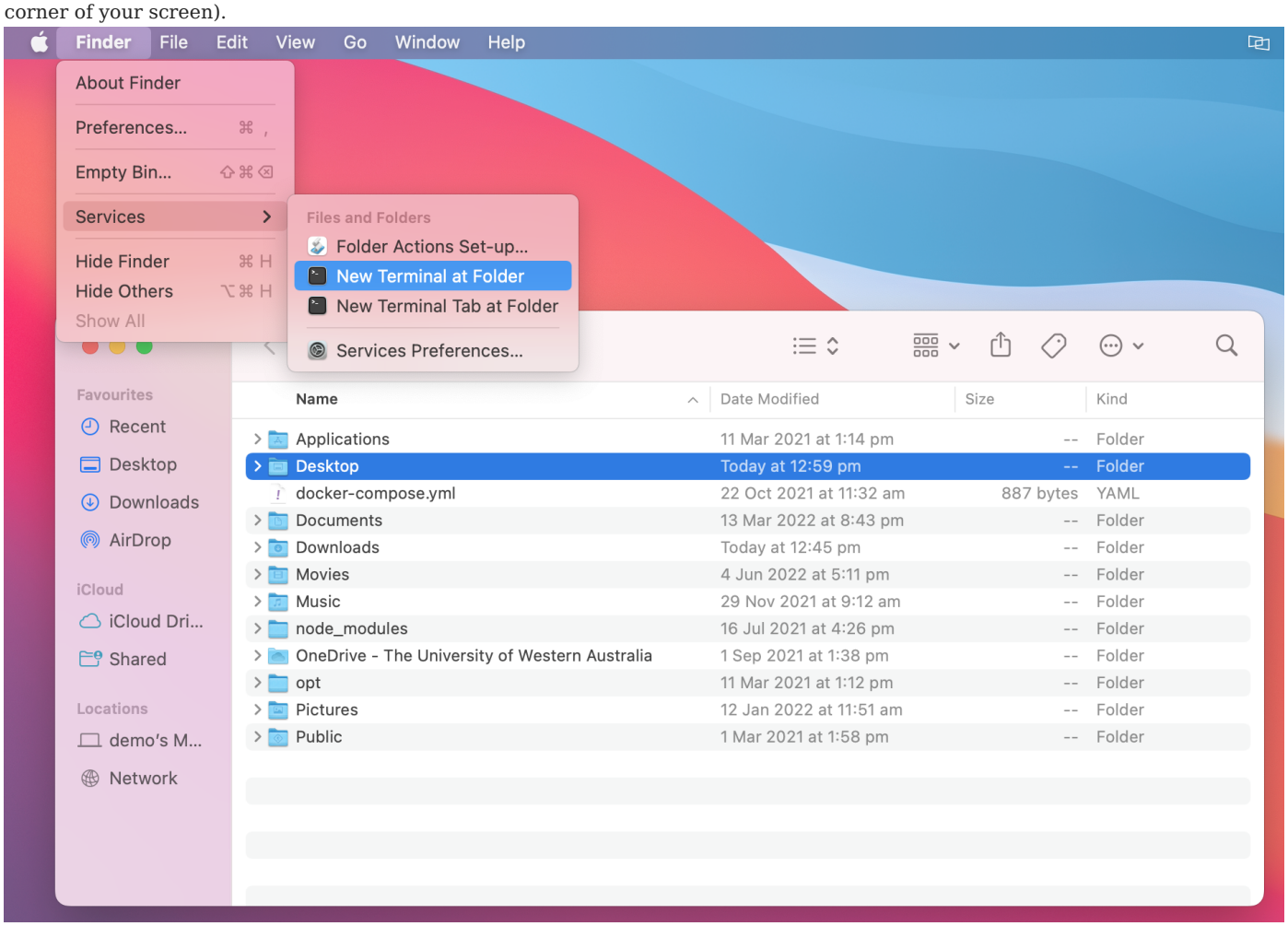
Students undertaking their cap-stone computer science, data science and software engineering unit, Professional Computing at the University of Western Australia, during semester 2 of 2022. Dominic Cain, Hannah Doret, Nathan Eden, Rachel Nguyen and Aidan Smith were responsible for the creation of this project.

2. The purpose of this documentation

This documentation is directed towards both the general users, and potential maintainers of this application. This documentation outlines the use and installation of this application, along with the packages/libraries included in the application.

3. Installation Guide

1. On your **MacOS** computer, use the **Finder** to **select the folder** you wish to download the application to.
2. Once the folder is selected, click **Finder**, choose **Services > New Terminal at Folder** (Located next to the Apple menu in the corner of your screen).

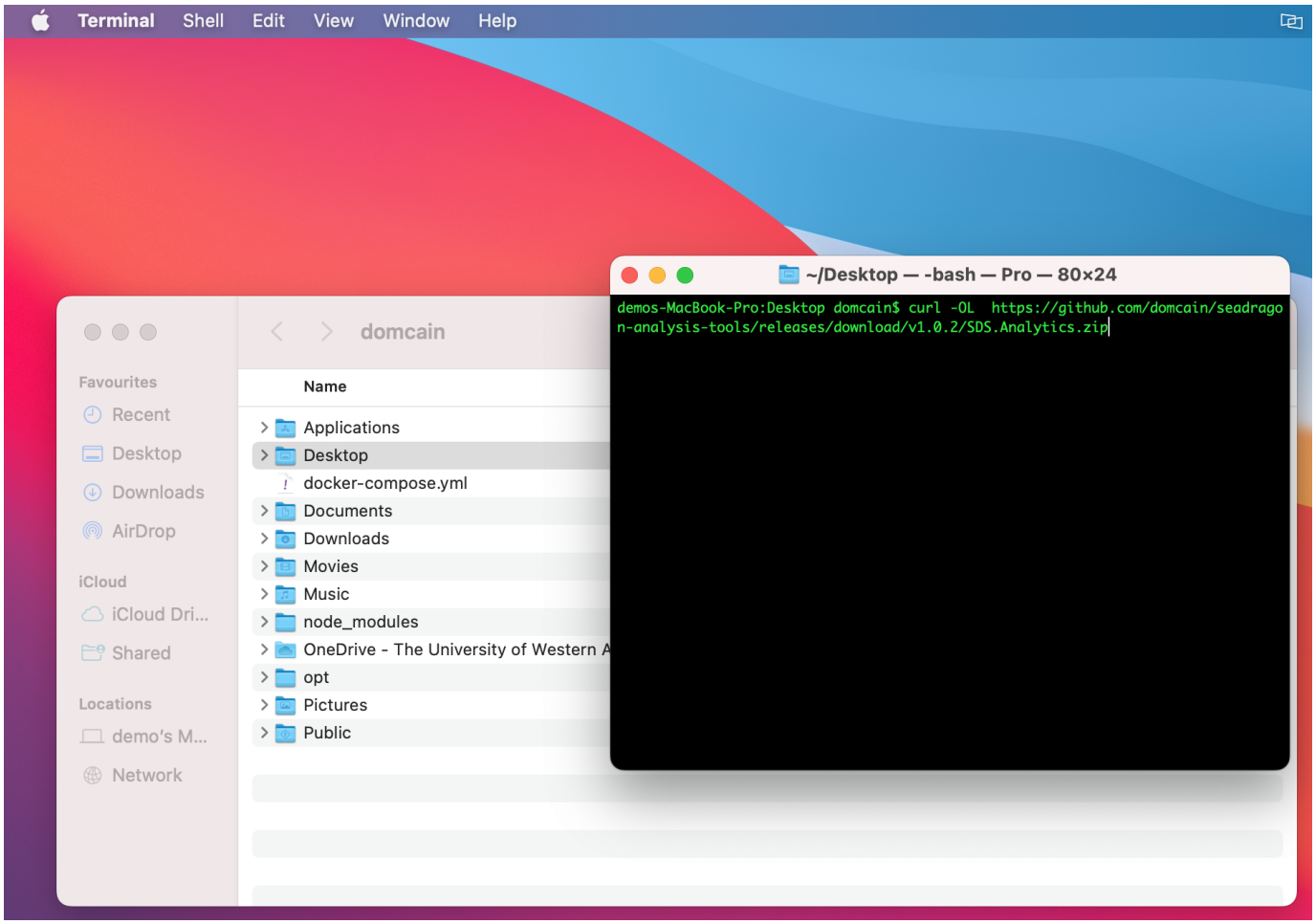


3. To download the project, **copy the following command**:

```
curl -OL  https://github.com/domcain/seadragon-analysis-tools/releases/download/v1.0.2/SDS.Analytics.zip
```

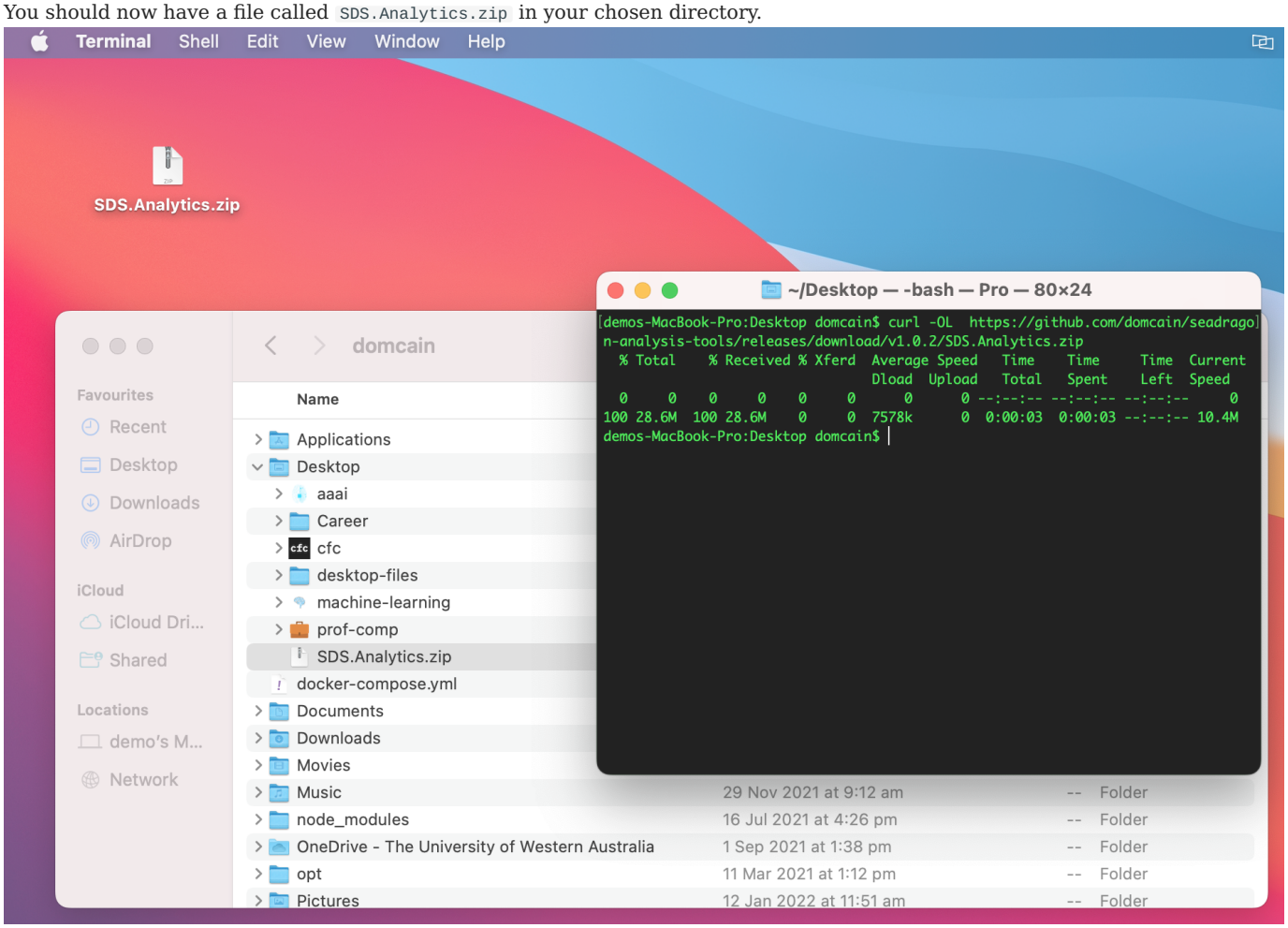
4. **Paste the command** into your command-line.

> **What does this do?**
>
> This accesses the website where the code is located and downloads the application. Performing this manually will result in bugs related to security interference from GitHub

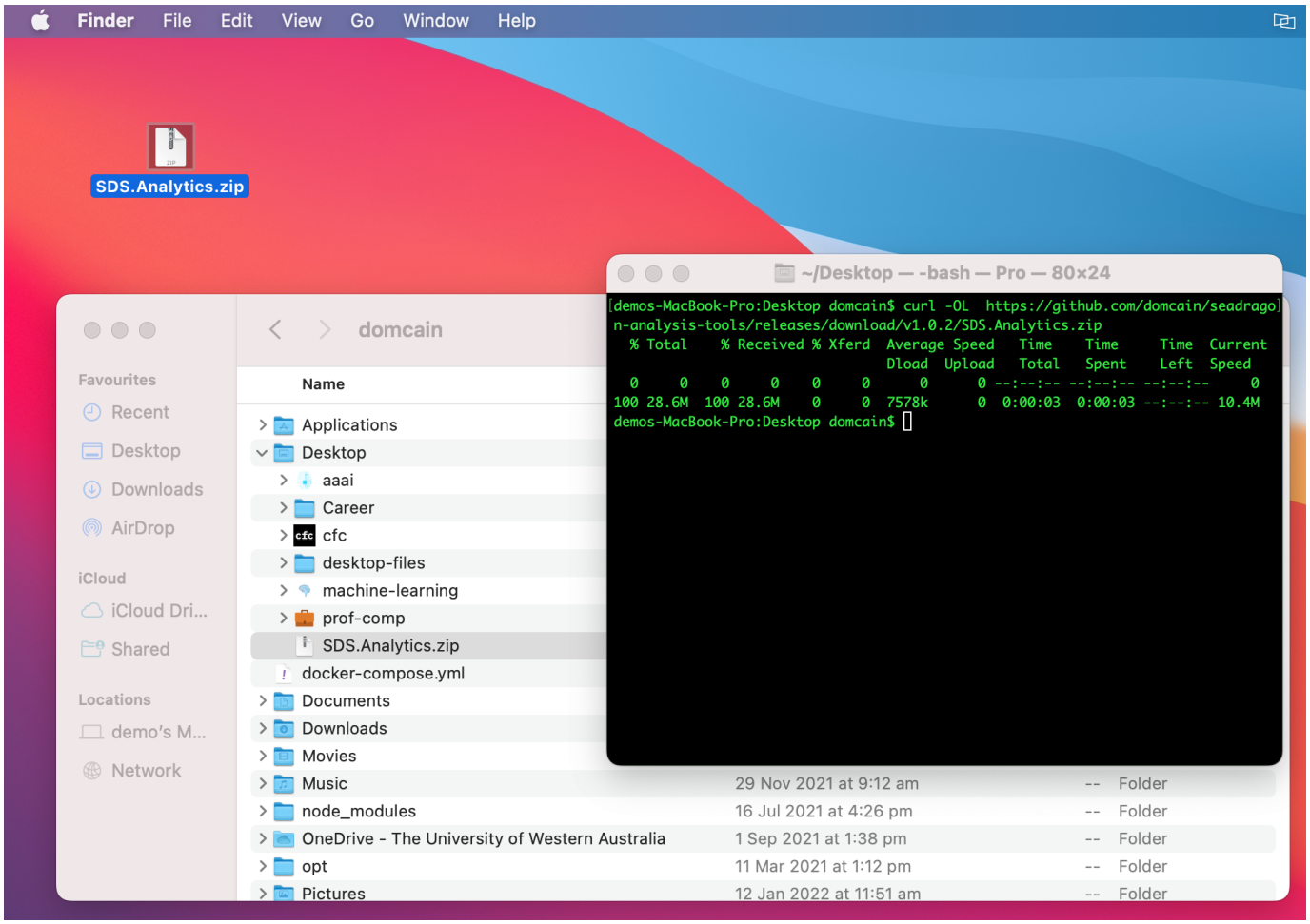


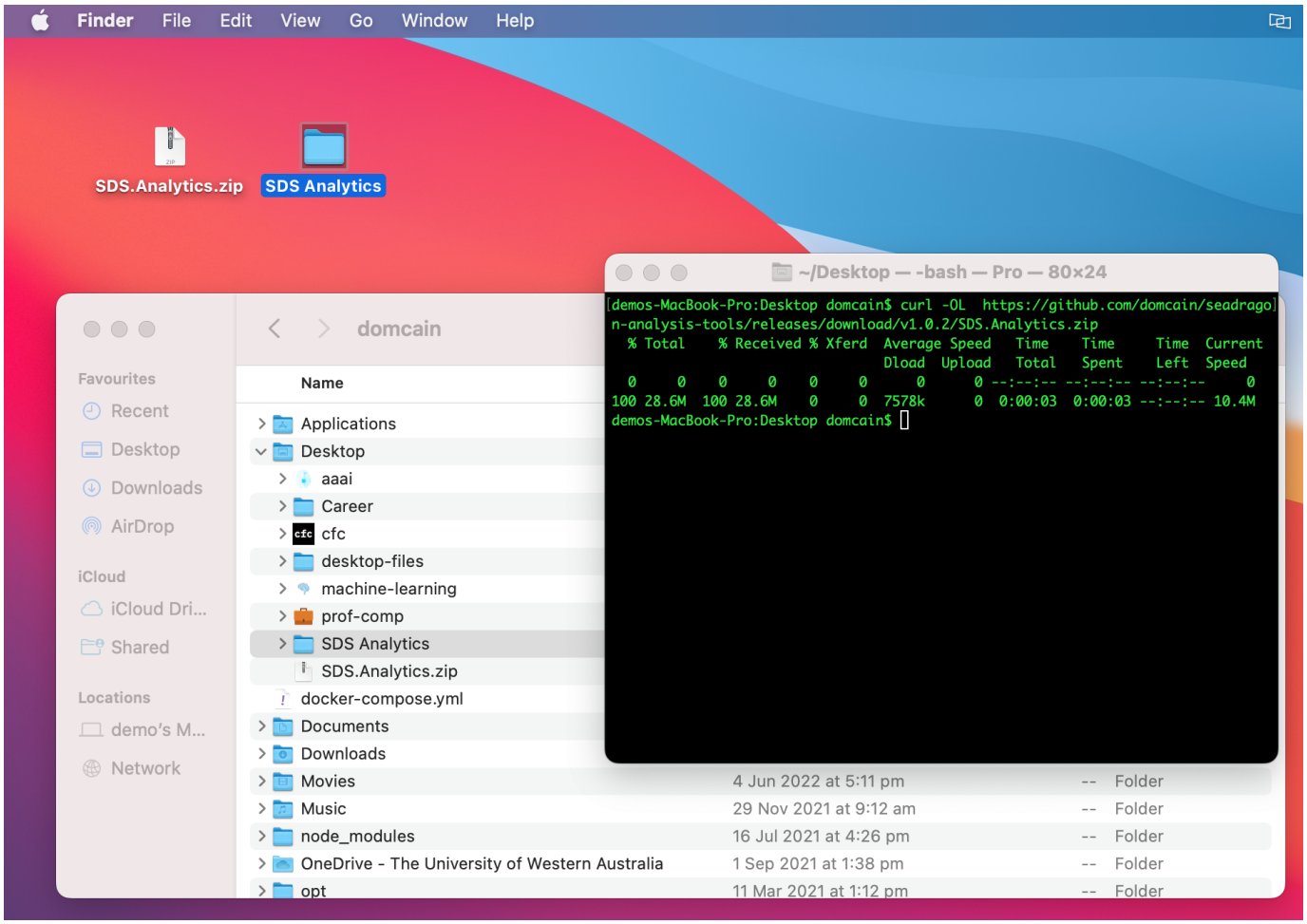
5. **Click enter** *if* the command has not already started.

You should now have a file called `SDS.Analytics.zip` in your chosen directory.



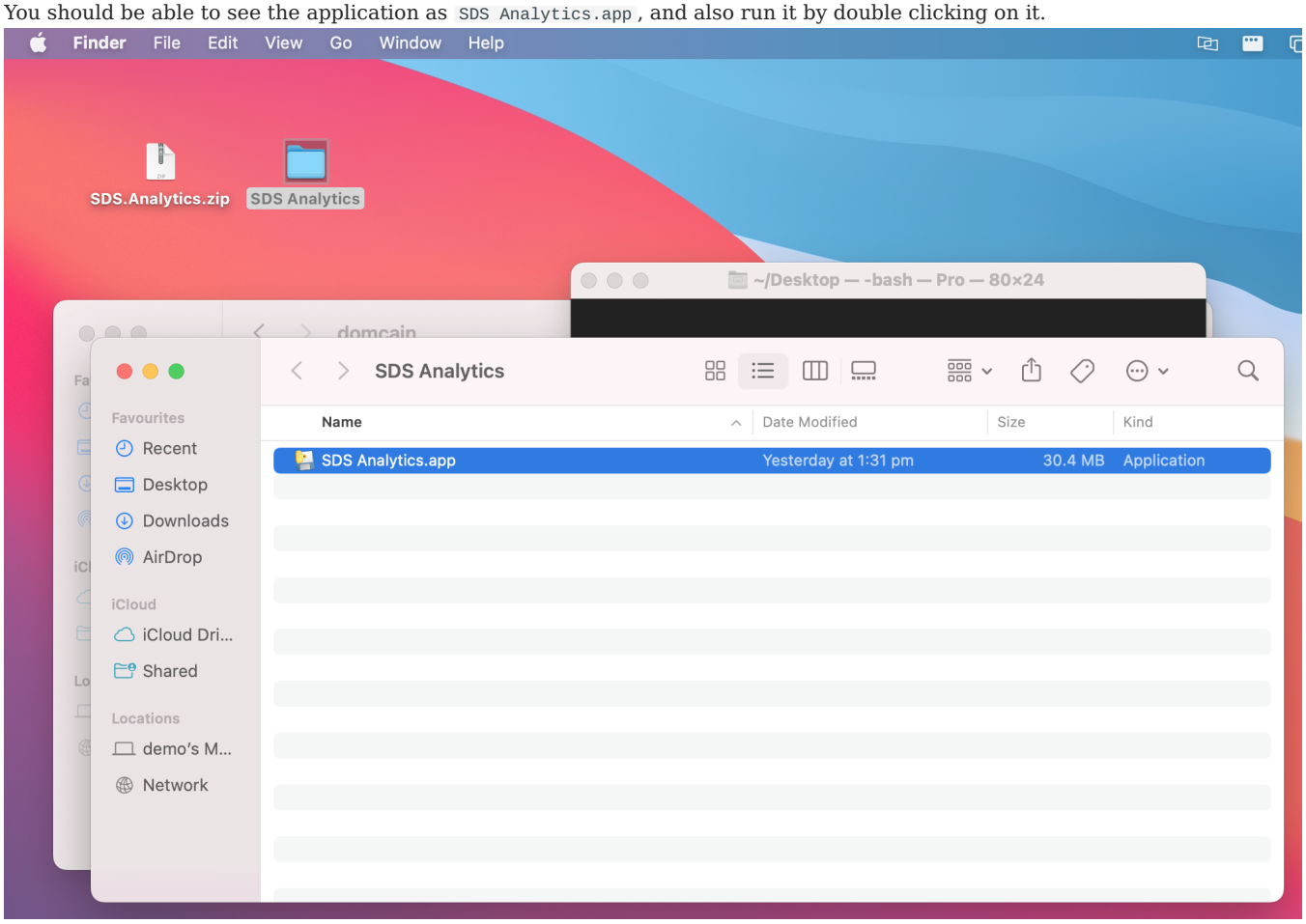
6. **Unzip** the `SDS.Analytics.zip` folder by double clicking on it with your mouse.


Finder File Edit View Go Window Help
SDS.Analytics.zip
~/Desktop — -bash — Pro — 80×24
[demos-MacBook-Pro:Desktop domcain$ curl -OL https://github.com/domcain/seadrago]
n-analysis-tools/releases/download/v1.0.2/SDS.Analytics.zip
% Total % Received % Xferd Average Speed Time Time Time Current
Dload Upload Total Spent Left Speed
0 0 0 0 0 0 0 0 --:--:-- --:--:-- --:--:-- 0
100 28.6M 100 28.6M 0 0 7578k 0 0:00:03 0:00:03 --:--:-- 10.4M
demos-MacBook-Pro:Desktop domcain$
domcain
Favourites
Recent
Desktop
Downloads
AirDrop
iCloud
iCloud Dri...
Shared
Locations
demo's M...
Network
Name
Applications
Desktop
aaai
Career
cfc
desktop-files
machine-learning
prof-comp
SDS.Analytics.zip
docker-compose.yml
Documents
Downloads
Movies
Music
node_modules
OneDrive - The University of Western Australia
opt
Pictures
29 Nov 2021 at 9:12 am -- Folder
16 Jul 2021 at 4:26 pm -- Folder
1 Sep 2021 at 1:38 pm -- Folder
11 Mar 2021 at 1:12 pm -- Folder
12 Jan 2022 at 11:51 am -- Folder



8. **Open** the `SDS Analytics` folder.

- You should be able to see the application as `SDS Analytics.app`, and also run it by double clicking on it.



Want to store the application elsewhere?

Go for it! The application can be run from anywhere on your computer!

4. User Manual

4.1 Upload SeadragonSearch file

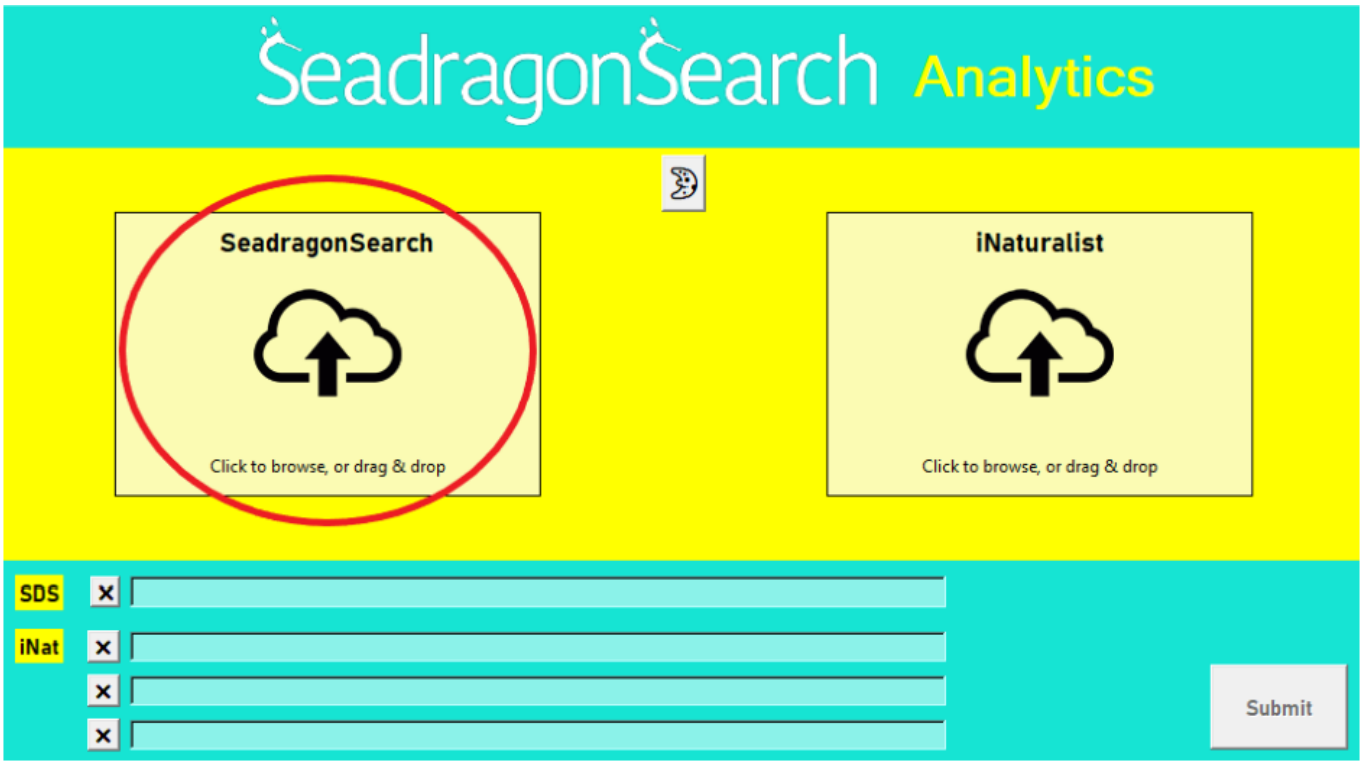
A maximum of 1 SeadragonSearch file can be uploaded at a time. The file needs to be an excel file. You can upload the file in two ways:

1. **'Click to browse'**

a. **Click** the SeadragonSearch upload box

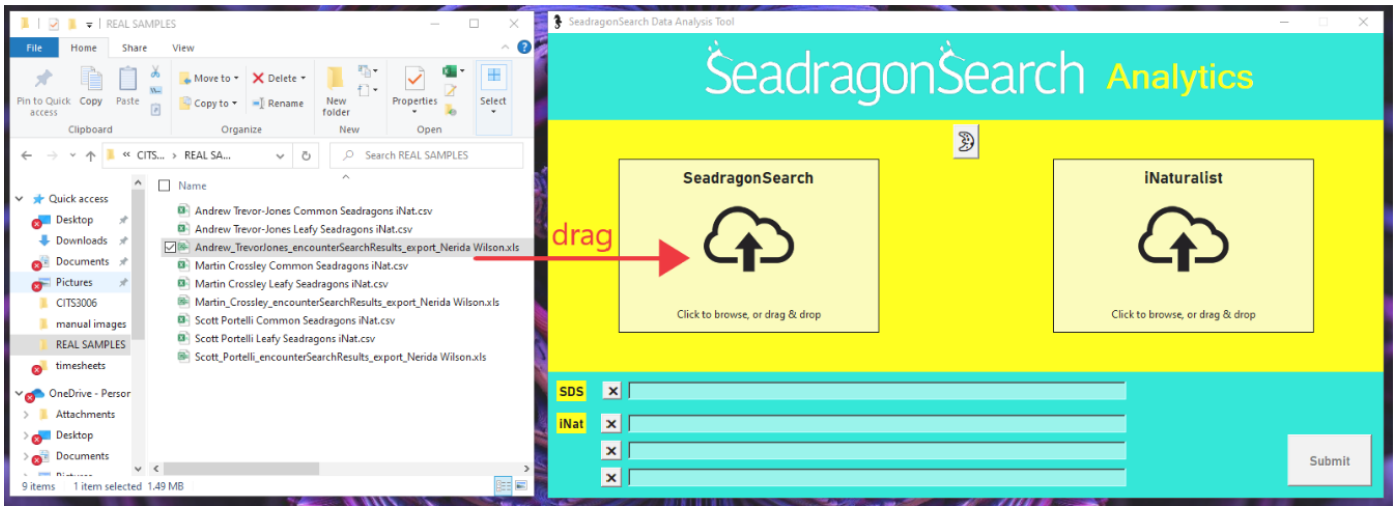
b. **Select** the SeadragonSearch excel file from the file explorer pop up.

2.



'Drag & drop'

From a separate file explorer window, **drag the SeadragonSearch excel file into the SeadragonSearch upload box**



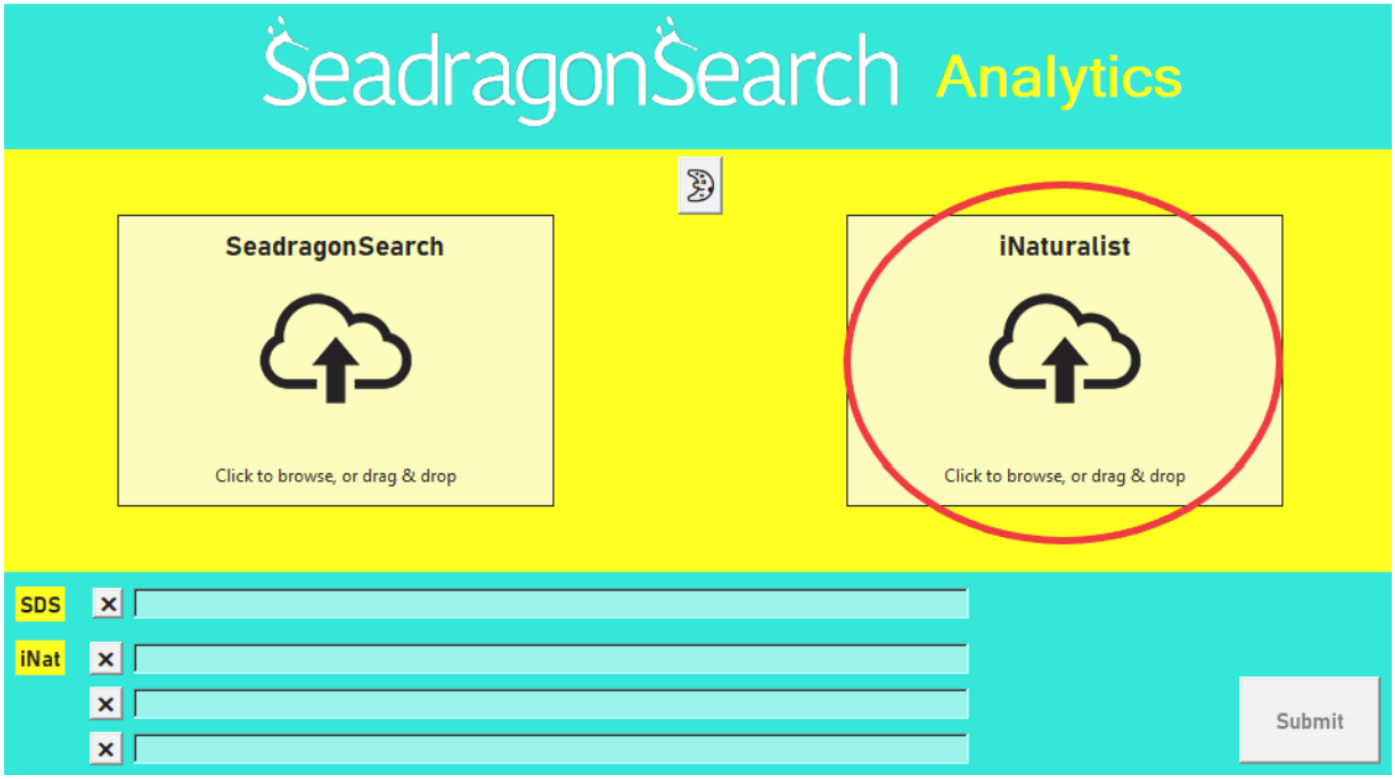
4.2 Upload iNaturalist file(s)

A maximum of 3 iNaturalist files can be uploaded at a time. This is to accommodate the Common, Leafy and Ruby Seadragon species data being stored in separate files. **Upload these files in any order**. iNaturalist files must be csv files. Supported extensions are **.csv** and **.txt**. Again, you can upload the file in two ways:

1. **'Click to browse'**

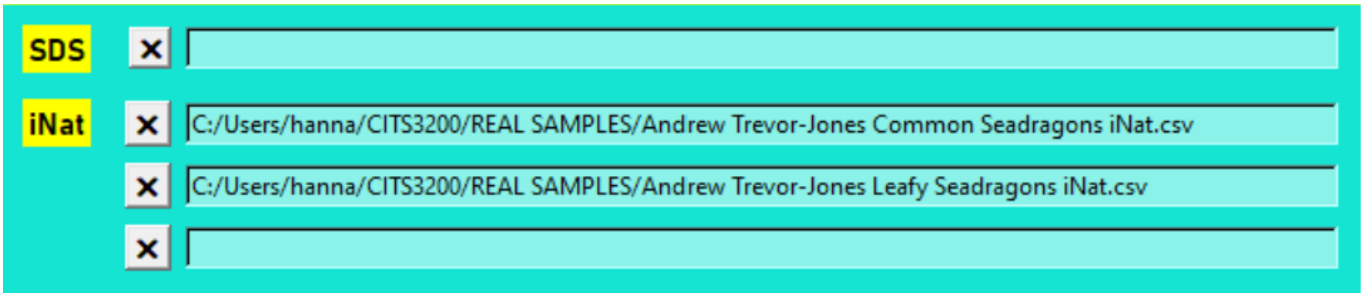
a. **Click** the iNaturalist upload box

b. **Select up to 3 iNaturalist .csv files** from the file explorer pop up



2. **'Drag & drop'**

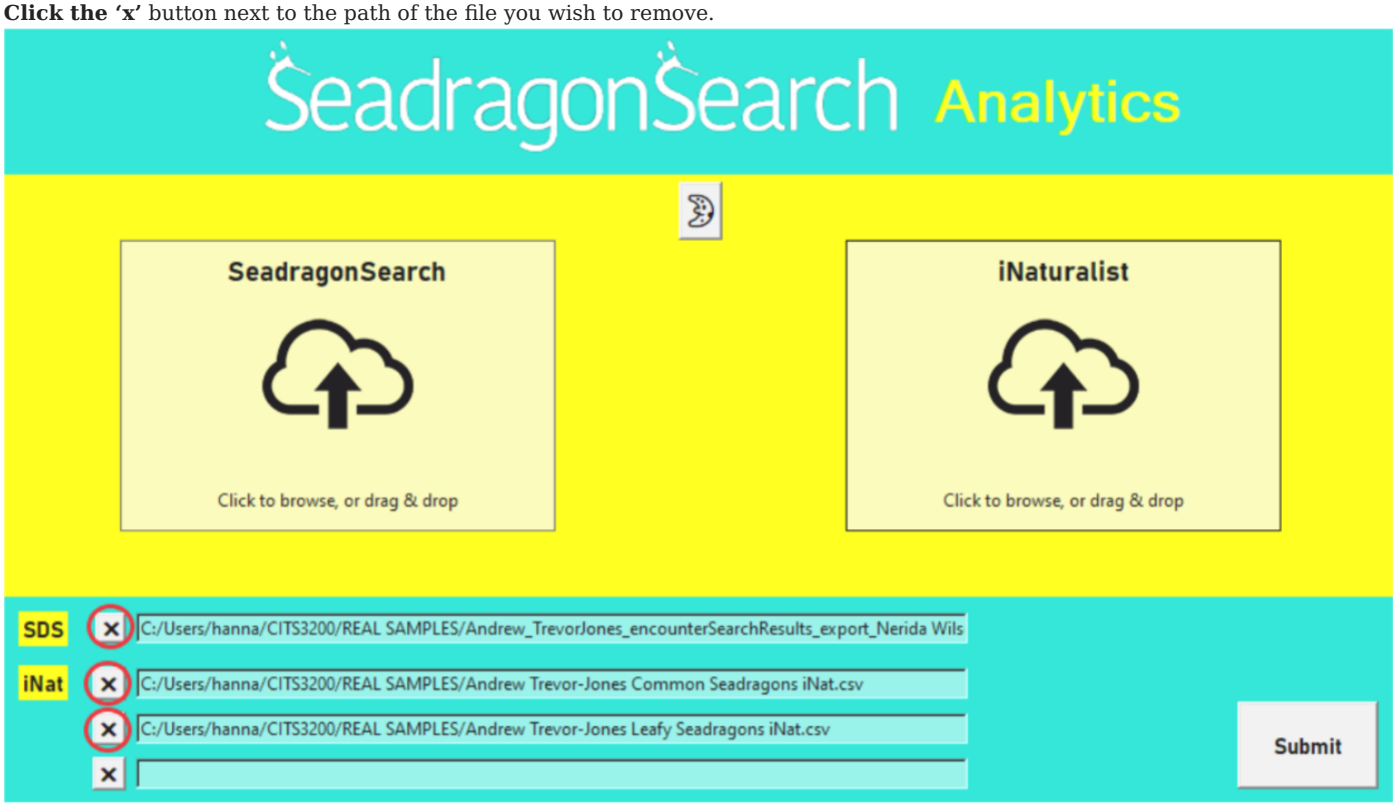
a. From a separate file explorer window, **drag up to 3 iNaturalist .csv files** into the iNaturalist upload box



4.3 Removing files

To remove any uploaded file:

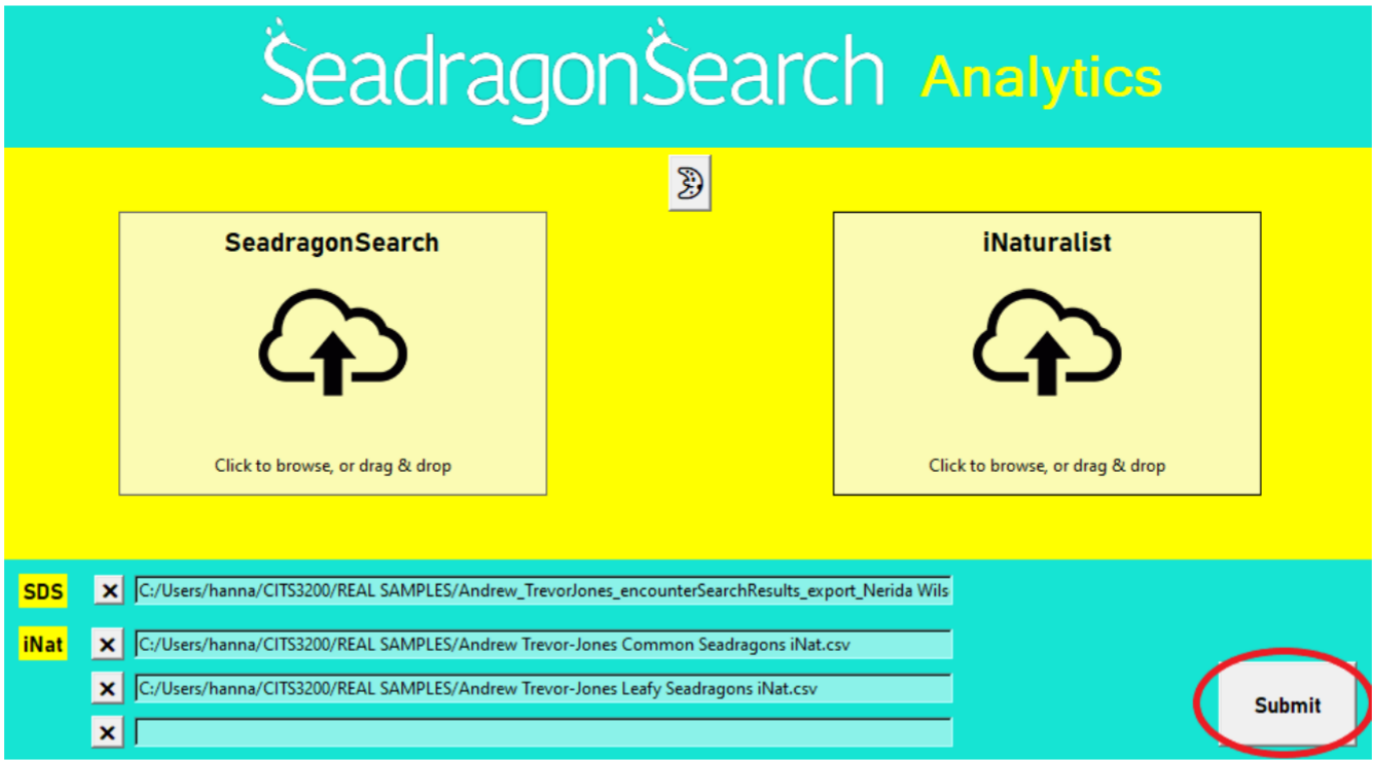
1. **Click the 'x'** button next to the path of the file you wish to remove.



4.4 Analysing the files

Once you are satisfied you have uploaded the files you wish to analyse, **click 'Submit'**.

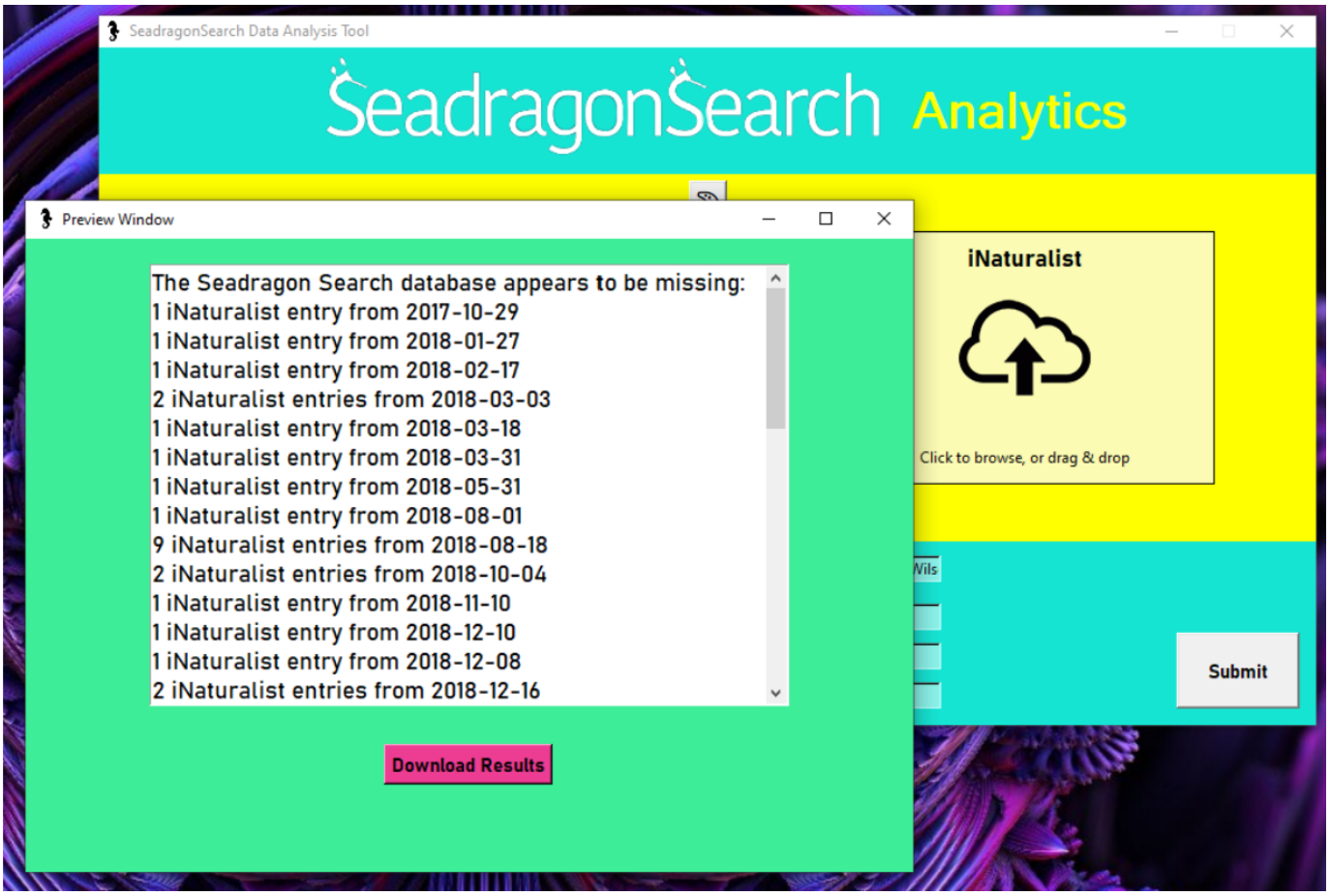
> **Note**
>
> The 'Submit' button will only be enabled when **at least 1 file** has been uploaded **for both SeadragonSearch and iNaturalist**.



Clicking ‘Submit’ will bring up a pop up window containing a preview of the results of the analysis.

Note

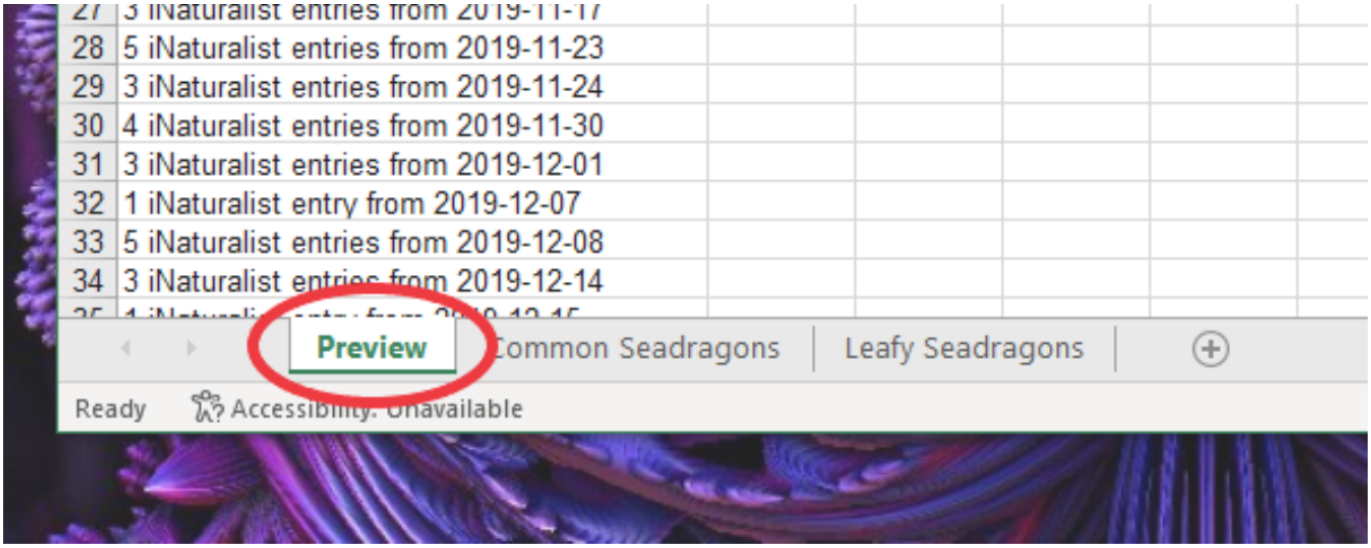
If you wish to download the full details of the results, **click ‘Download Results’**.



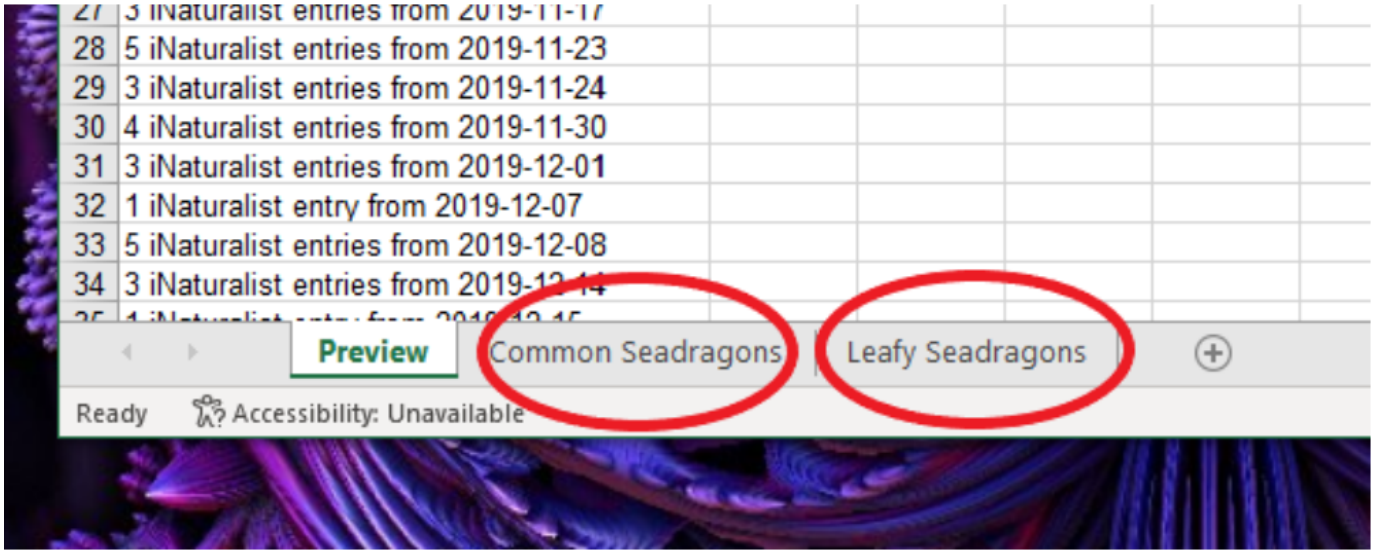
From the file explorer pop up, you can navigate to where you would like the results file to be downloaded, as well as rename the file to something of your choice.

4.5 Interpreting the results file

The results file will be an Excel (.xls) document. The first worksheet in the file is a copy of the preview that was provided when the results file was generated.



Each subsequent worksheet in the file corresponds to a specific iNaturalist file which was uploaded. For example, if you uploaded an iNaturalist file for both Common and Leafy Seadragons, the results file will have a worksheet for each of these species, as in the example below.



These worksheets are copies of the iNaturalist files which were uploaded for analysis, with highlighted rows to indicate entries that may be missing from the SeadragonSearch database. Each highlighted entry should be reviewed by a human before being added to the SeadragonSearch database.

5. Running the code in the development environment

1. **Clone the repository** to your personal computer by running:

```
$   git clone https://github.com/domcain/seadragon-analysis-tools.git
```

   in your terminal.
2. In your terminal, **navigate to/open the cloned repository**.
3. From the seadragon-analysis-tools/ directory, **install the required dependencies** to run the application locally.

- You can do this using the command:
- MacOS:

```
pip install -r requirements.txt
```

- Windows:

```
py -m pip install -r requirements.txt
```

4. From the `/src` directory, run:

```
python3 'SDS Analytics.py'
```

6. Project Layout

```
seadragon-analysis-tools/
    docs/
        mkdocs.yml                  # The documentation pages configuration file.
        requirements.txt            # Dependencies required to run the documentation locally.
        docs/
            index.md                # The documentation homepage.
            for-the-developer.md    # Documentation relevent to code maintainers.
            user-manual.md          # Documentation relevent to users of the application.
            packages-used.md        #
            images/
                seadragon.png       # Icon in top left of documentation page.
    src/
        SDS Analytics.py            # Main UI code.
        data_analysis.py            # Code for producing output files.
        images/
           cloud.png                # Image used in the UI.
           sdstitle.png             # Image used in the UI.
           seahorse.gif             # Special image for MacOS Dock icon.
    Test_case/
        Test_cases.csv              #
        test_data_analysis.py       # Code for testing the output file.
        1/
            inat1.csv               # Test 1.
        2/
            inat2.csv               # Test 2.
        3/
            inat3.csv               # Test 3.
        4/
            inat4.csv               # Test 4.
        5/
            inat5.csv               # Test 5.
        6/
            inat6.csv               # Test 6.
        7/
            inat7.csv               # Test 7.
        8/
            inat8.csv               # Test 8.
        9/
            inat9.csv               # Test 9.
        10/
            inat10.csv              # Test 10.
        11/
            inat11.csv              # Test 11.
        12/
            inat12.csv              # Test 12.
        13/
            inat13.csv              # Test 13.
        14/
            inat14.csv              # Test 14.
    .gitignore                      # File to minimise unnessesary repository file contributions.
    hook-tkinterdnd2.py             # Hook file required to build a python application that uses tkinterdnd2.
    LICENSE.txt                     # License (Creative Commons).
    README.md                       # Initial documentation.
    requirements.txt                # Dependencies required to run the application locally.
```

6.1 SDS Analytics.py

`SDS Analytics.py` acts as the `main.py` seen in a variety of other projects. This file creates the user interface and is responsible for all its functionality.

6.2 data_analysis.py

`data_analysis.py` handles all under the hood operations to produce an Excel workbook from input filepaths, by comparing the contents of the input files.

The only function in `data_analysis.py` that should be called from another file is called analyse_data_files. This function takes two arguments: the first is the filepath of a Seadragon Search file (string), and the second is a list of the filepaths of iNaturalist files (a list of strings). If the function encounters an error, it returns a list of the form [False, error_message] (where error_message is a string). If the function executes successfully, it returns a list of the form [True, preview, suggested_filename, new_wb] (where preview is a string, suggested_filename is a string, and new_wb is a Workbook, which is a type from the module xlwt).

7. Packages Used

Listed below are the imported python modules used in this project.

Need more info?

Click on any of the headings to read module documentation.

7.1 User Interface

`SDS Analytics.py` acts as the `main.py` seen in a variety of other projects. This file creates the user interface.

tkinter is the standard Python interface to the Tcl/Tk GUI toolkit. This package provides the building blocks of the user interface.

tkinterdnd2: This package provides the 'drag & drop' functionality to the interface.

7.2 Data Analysis

`data_analysis.py` handles all under the hood operations taking input file paths, and producing an excel file comparing the contents of the input files.

xlrd: This package provides the tools required to read and analyse data from input excel files.

xlwt: This package provides the tools to generate the output excel file.

openpyxl: This package provides the tools to read/write excel files other than .xls, such as .xlsx, .xlsm, .xltx, .xltm