For this project, you’ll act as a data analyst for the National Park Service. You’ll be helping them analyze data on endangered species from several different parks.

The National Parks Service would like you to perform some data analysis on the conservation statuses of these species and to investigate if there are any patterns or themes to the types of species that become endangered. During this project, you will analyze, clean up, and plot data, pose questions and seek to answer them in a meaningful way.

After you perform analysis, you’ll be creating a presentation to share your findings with the National Park Service.

Off-Platform vs. On-Platform

There are two possible ways of completing this project:

You can complete this project on your own computer. To do this, you’ll need to install Python and Jupyter notebooks, as well as use the command line. We recommend this option, because it will help prepare you for pursuing more data analysis projects on your own.

You can complete the project on Codecademy.com. You won’t need to install Python and you won’t need to do anything on the Command Line. If you want to complete the project online, you can find it here: Biodiversity in National Parks: Capstone Project

Task 1: Analyzing Data

Download biodiversity.zip

Double click on it to “unzip” the folder. It should contain several items:

observations.csv, species\_info.csv, biodiversity.ipynb, biodiversity-solutions.ipynb 3. In the command line, navigate into the biodiversity directory 4. Type the following into the command line:

jupyter notebook

This should open a browser tab. Click on biodiversity.ipynb in the browser tab. This will open up your Jupyter Notebook.

Follow the steps in the Jupyter Notebook. They will help guide you through your analysis.

If you ever get stuck during this project, feel free to open the biodiversity-solutions.ipynb. Warning: it contains all of the solution code for the project, so only use it if you are really stuck!

Task 2: Create a Slide Deck

Once you’ve performed your analysis, either on your computer or on Codecademy, you’re ready to create your slide deck. One of the most important parts of Data Analysis is being able to communicate your findings.

Create a slide deck using Google Drive, Microsoft Powerpoint, or some other presentation software. Your presentation should include the following:

A title slide

A section describing the data in species\_info.csv. Be sure to include some (or all) of what you noticed while working through the notebook.

A section describing the significance calculations that you did for endangered status between different categories of species.

A recommendation for conservationists concerned about endangered species, based on your significance calculations

A section describing the sample size determination that you did for the foot and mouth disease study

All of the graphs that you created in the notebook

If you like, you can also record a video of yourself giving the presentation and upload it to YouTube.

Task 3: Share your Presentation!

Now that you have completed this capstone, you can share you presentation! We’ll give you instructions on how to share your project after the next capstone project.