

## Cogniac Coding Challenge (Python)

### Challenge:

First, retrieve the compressed tar archive, [https://s3-us-west-1.amazonaws.com/cogniac-public-data/python\\_coding\\_challenge.tar.gz](https://s3-us-west-1.amazonaws.com/cogniac-public-data/python_coding_challenge.tar.gz), which contains the following:

1. modified\_images/
2. score.py
3. validation\_set.csv

We have gathered various images from Flickr <https://www.flickr.com/>. These images were then modified in various ways (resized, rotated, contrast changes, etc.) to create a modified image dataset (“modified\_images/”). Using the Flickr API, search Flickr for images with the tag “people” and user id “143060054@N02”. Write a Python script that matches the modified images to the Flickr original from which it was modified. There are many images in the dataset with modifications of varying difficulty and type. Try to match as many as you can. You can use any third party libraries you like but be prepared to explain the concepts behind any code you use whether you wrote it or not. In the end, we must be able to run your code so that it repeats your results.

If needed, you can use the following Flickr API key and secret:

**Key:** c871b759c44a9c7d3ced100ba2cf4dd0

**Secret:** da8022494490182e

Your code should output a CSV file called “results.csv” which lists the file name of the modified images to the URL of the original images on Flickr. Look at the (filename, URL) formatting of “validation\_set.csv” for a reference. validation\_set.csv contains the correct original URL’s for several of the modified images. You can use score.py to test your “results.csv” file with the following usage:

```
python score.py results.csv
```

### Deliverables:

We would like you to create a public GitHub repository to work in and send us a link to it. We prefer that you commit often enough to allow us to view and understand the progression of your code. Your code should output a CSV file called “results.csv” which lists the file name of the modified images to the URL of the original images on Flickr. Look at the (filename, URL) formatting of “validation\_set.csv” for a reference. In the end, your GitHub repository should contain the following:

1. Your python code
2. The CSV file of your results labeled “results.csv”.
3. Make a brief write up (a page or two) of the approach you took to complete this challenge, any metrics that you find relevant to describe your results, and how you could improve your code given more time.

If you have any questions don’t hesitate to email [matt@cogniac.co](mailto:matt@cogniac.co)