

## Data Appendix

This appendix contains information about the dataset and its variables. The data file used for this project consists of a collection of labeled snake images and associated annotation files.

*Snake Image Data.* This dataset contains images of common snake species found in the DMV area. It was created and published by Chris Workspace on Roboflow, with image sources from iNaturalist. The dataset includes 889 total images (823 unique), divided into four species—two venomous (Timber Rattlesnake and Eastern Copperhead) and two non-venomous (Common Garter and Black Rat). Each image is annotated with bounding box metadata provided in JSON format. The data was pre-split into training (90%) and test (10%) sets.

### Data Dictionary

Variable Name	Description
images	Image name
label	Species label
x	x coordinate
y	y coordinate
width	Width of the snake
height	Height of the snake

### Exploratory Plots

```
Train:
- Total annotations: 806
- Unique images: 749

Validation:
- Total annotations: 83
- Unique images: 74
```

