

## **DS4A Colombia AI Training:**

### **Final Project Example Datasets**

Aggregations of dataset sources:

- <https://www.kdnuggets.com/datasets/index.html>
- <https://github.com/awesomedata/awesome-public-datasets>
- <https://www.kaggle.com/>
- <https://www.datasetlist.com/>
- <https://blog.cambridgespark.com/50-free-machine-learning-datasets-image-datasets-241852b03b49>

Aggregations of dataset sources (sign up required):

- <https://community.ibm.com/community/user/businessanalytics/home?content-type=sample-data#resource-table>
- <https://data.world/>

Open datasets, lots of medical:

- <http://archive.ics.uci.edu/ml/index.php>

Dataset search:

- <https://www.google.com/publicdata/directory>
- <https://registry.opendata.aws/>

Open datasets - governments:

- <https://data.worldbank.org/>
- <https://www.who.int/gho/database/en/>
- <https://data.europa.eu/euodp/en/data/>

Example Datasets:

- <https://www.kaggle.com/panchicore/vehicles-colombia-fasecolda>
- <https://www.kaggle.com/ruslankl/european-union-lgbt-survey-2012>
- <https://www.kaggle.com/samhiatt/xenocanto-avian-vocalizations-canv-usa>
- <http://pascal.inrialpes.fr/data/human/>