**Datasets - repositories**

* [Kaggle](https://www.kaggle.com/)
  + [UCI](http://archive.ics.uci.edu/ml/index.php) data repository (also at kaggle) - mostly older and smaller
* [Lionbridge](https://lionbridge.ai/datasets/) - lots by language and application
  + [top 50](https://lionbridge.ai/datasets/the-50-best-free-datasets-for-machine-learning/)
  + [econ/government](https://lionbridge.ai/datasets/17-best-finance-economic-datasets-for-machine-learning/)
* <https://datausa.io/>
* <https://towardsdatascience.com/top-sources-for-machine-learning-datasets-bb6d0dc3378b>
* <https://stats.idre.ucla.edu/other/dae/> - regression examples
* <https://biodatamining.biomedcentral.com/articles/10.1186/s13040-017-0154-4> - bio-related
* [R dataset collection](https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/00Index.html)
* Cambridge Spark
  + [finance and econ](https://blog.cambridgespark.com/50-free-machine-learning-datasets-part-two-financial-and-economic-datasets-6620274ee593)
  + [images](https://blog.cambridgespark.com/50-free-machine-learning-datasets-image-datasets-241852b03b49)
  + [sentiment analysis](https://blog.cambridgespark.com/50-free-machine-learning-datasets-sentiment-analysis-b9388f79c124)
  + [government](https://blog.cambridgespark.com/50-free-machine-learning-datasets-part-one-government-data-portals-e39524ba601b)
* [Physionet Challenges](https://physionet.org/challenge/): physiological time series classification

**Datasets - examples**

* <http://archive.ics.uci.edu/ml/datasets/Heart+Disease> - Cleveland heart disease (classification classic)
* [Boston housing](https://www.kaggle.com/c/boston-housing) (regression classic)
  + Explored in <https://towardsdatascience.com/linear-regression-on-boston-housing-dataset-f409b7e4a155>
* [Wine quality](https://www.kaggle.com/rajyellow46/wine-quality) (classification - red/white; regression - quality)
  + Explored in <https://www.datacamp.com/community/tutorials/deep-learning-python#finetune>
* [wine reviews](https://www.kaggle.com/zynicide/wine-reviews)
* [yelp reviews](https://www.yelp.com/dataset)
* [movie reviews](https://www.kaggle.com/c/sentiment-analysis-on-movie-reviews)

**RL simulations**

* <https://github.com/kengz/awesome-deep-rl> - excellent collection!
* [Inverted pendulum](http://ctms.engin.umich.edu/CTMS/index.php?example=InvertedPendulum&section=SimulinkModeling) - classic toy problem
  + And the Mountain Car (Sutton and Barto)