

eDatasets - repositories

- [Kaggle](#)
 - [UCI](#) data repository (also at kaggle) - mostly older and smaller
- [Lionbridge](#) - lots by language and application
 - [top 50](#)
 - [econ/government](#)
- <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](#)
- Cambridge Spark
 - [finance and econ](#)
 - [images](#)
 - [sentiment analysis](#)
 - [government](#)
- [Physionet Challenges](#): physiological time series classification

Datasets - examples

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

RL simulations

- <https://github.com/kengz/awesome-deep-rl> - excellent collection!
- [Inverted pendulum](#) - classic toy problem
 - And the Mountain Car (Sutton and Barto)