

NSF supported WSSU Faculty Workshop on Data Science Pedagogy & Practice

Useful Links

Excel

- Data Analysis with Excel, <https://www.excel-easy.com/data-analysis.html>
- Excel Pivot Table, <https://www.dummies.com/software/microsoft-office/excel/excel-data-analysis-creating-an-excel-pivot-table/>

Tableau

- Getting Started with Tableau, <http://onlinehelp.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-home.html>

Python

- Great Interactive Way to Learn Python Online, <http://www.learnpython.org/>
- Learn Python in 10 minutes (Cheatsheet), <https://www.stavros.io/tutorials/python/>
- Python Official Documentation, <https://docs.python.org/2/tutorial/>

Jupyter Notebook

- Installing Jupyter using Anaconda (Python 2.7), <https://jupyter.org/install>

Python Visualization

- Matplotlib Tutorial, https://matplotlib.org/users/pyplot_tutorial.html

Python Panda

- Panda Intro, http://pandas.pydata.org/pandas-docs/stable/getting_started/10min.html
- Panda vs. SQL, http://pandas.pydata.org/pandas-docs/stable/getting_started/comparison/comparison_with_sql.html

R Language

- R Tutorial, <http://www.cyclismo.org/tutorial/R/>
- Quick-R by DataCamp, <https://www.statmethods.net>
- Python vs. R, <https://www.datacamp.com/community/tutorials/r-or-python-for-data-analysis#gs.nrBsDZQ>

Machine Learning

- Intro to ML - coursera, <https://www.coursera.org/lecture/machine-learning/what-is-machine-learning-Ujm7v>

Regression Tutorials

- Introduction to Linear Regression, <http://onlinestatbook.com/2/regression/intro.html>
- Correlation and linear regression, <http://www.biostathandbook.com/linearregression.html>

Classification

- K Nearest Neighbors - Classification, http://www.saedsayad.com/k_nearest_neighbors.htm
- Decision Trees, http://www.saedsayad.com/decision_tree.htm

- Naive Bayesian, http://www.saedsayad.com/naive_bayesian.htm

Clustering

- K-Means Clustering, http://www.saedsayad.com/clustering_kmeans.htm

Machine Learning with Python

- Applied Machine Learning in Python Course, <https://www.coursera.org/learn/python-machine-learning>
- Scikit-Learn Tutorial + Documentation, <http://scikit-learn.org/stable/tutorial/basic/tutorial.html>

Machine Learning with Weka

- Install Weka, <https://www.cs.waikato.ac.nz/ml/weka/downloading.html>
- Weka Documentation, <https://www.cs.waikato.ac.nz/ml/weka/documentation.html>
- Basic Weka Tutorial, <https://machinelearningmastery.com/how-to-run-your-first-classifier-in-weka/>

DataSets

- Weka Datasets, <https://www.cs.waikato.ac.nz/ml/weka/datasets.html>
- Kaggle Dataset, <https://www.kaggle.com/datasets>
- UCI ML Repository, <https://archive.ics.uci.edu/ml/datasets.php>
- Statistical Abstracts of the US (census), http://www.census.gov/library/publications/time-series/statistical_abstracts.html
- World Bank Data, <http://data.worldbank.org/>
- FiveThirtyEight, <https://github.com/fivethirtyeight/data>
- US Government's Open Data (data.gov), <http://www.data.gov/>
- Pew Research Center, <http://www.pewresearch.org/data/>
- The Data and Story Library, <http://lib.stat.cmu.edu/DASL/>
- Dr. John Rasp's Statistics Website, <http://www2.stetson.edu/~jrasp/data.htm>
- Awesome Public Datasets, <https://github.com/caesar0301/awesome-public-datasets>
- Cool Datasets, <https://www.cooldatasets.com/>

NSF Funding Opportunities

- The Proposal and Award Policies and Procedures Guide NSF 18-1, which contains the Grant Proposal Guide, is available at: https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf18001
- Harnessing the Data Revolution (HDR): Institutes for Data-Intensive Research in Science and Engineering - Ideas Labs (I-DIRSE-IL), https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505614
- Harnessing the Data Revolution (HDR): Institutes for Data-Intensive Research in Science and Engineering - Frameworks (I-DIRSE-FW), https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505631
- Harnessing the Data Revolution: Transdisciplinary Research in Principles of Data Science Phase I, https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505347
- Harnessing the Data Revolution (HDR): Data Science Corps (DSC), https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505536
- Inclusion across the Nation of Communities of Learners of Underrepresented Discoverers in Engineering and Science (INCLUDES), https://www.nsf.gov/publications/pub_summ.jsp?WT.z_pims_id=505630&ods_key=nsf19546
- HBCU-UP https://www.nsf.gov/publications/pub_summ.jsp?WT.z_pims_id=5481&ods_key=nsf18522

- Computational and Data-Enabled Science and Engineering (CDS&E)
https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504813&org=NSF&sel_org=NSF&from=fund
- Improving Undergraduate STEM Education: Education and Human Resources (IUSE: EHR)
https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505082&org=NSF&sel_org=NSF&from=fund
- Training-based Workforce Development for Advanced Cyberinfrastructure (CyberTraining)
https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505342&org=NSF&sel_org=OAC&from=fund
- Computer Science for All (CSforAll:RPP)
https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505359&org=NSF&sel_org=OAC&from=fund
- Improving Undergraduate STEM Education: Computing in Undergraduate Education (IUSE: CUE)
https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505630&org=NSF&sel_org=NSF&from=fund