DOE Project Resources

Project Files
Code Examples · Learning Resources · Other Examples · Past Sensors

▼ Table of Contents

Data Workflow

Data workflow adapted from Biozentrum University of Basel, Switzerland.

Feel free to use this as a reference when planning how to manage data on the DOE project.

Built With

All of my previous analysis was completed using Python and Jupyter Notebooks. Feel free to use any IDE you prefer. Google Collab could be a good option for working on analysis as a group. * **Python** * **Jupyter Notebook** * **Pandas** * **Matplotlib** * **Seaborn** * **Statsmodels** * **SciPy**

Other software that may may be useful depending on preferences. * **IBM SPSS** * **SAS/STATS** * **Excel Analysis ToolPak**

Useful Methods

These are just a few of my thoughts on some useful methods you may come across.

For testing the difference between groups: * ANOVA * Tukey's HSD * Boxplots

For time series analysis: * Linear and Multivariate Regression * Moving Averages * Pandas Groupby Function * Box-Jenkins ARIMA models * Box-Jenkins Multivariate Models * Holt-Winters Method

Basic Statistics: * Mean and Median * Standard Deviation

Common Visualizations: * Boxplot * Histogram * Scatterplot * Heatmap

Online Resources

Some pages and websites that you may find useful.

Working with time-series analysis in Python: * Time Series in Python — Exponential Smoothing and ARIMA processes * Time Series in Python — Part 2: Dealing with seasonal data * Time Series Decomposition In Python * Linear Regression in Python * Pandas GroupBy: Your Guide to Grouping Data in Python * A Guide to Time Series Analysis in Python * Time Series Analysis in Python – A Comprehensive Guide with Examples

Basic statistics in Python: * Python Statistics Fundamentals: How to Describe Your Data * How to Explain Data using Gaussian Distribution and Summary Statistics with Python * Descriptive Statistics with Python * Hands On Bayesian Statistics with Python, PyMC3 & ArviZ

ANOVA and Tukey's in Python: * How to Perform Tukey's Test in Python * ANOVA + Tukey Test In Python * N-Way ANOVA

Visualizations in Matplotlib and Seaborn: * Visualization with Matplotlib * Matplotlib Gallery * Python Graph Gallery

For more examples in Python, please refer to common sites such as: * Real Python * W3Schools * Python Data Science Handbook * Python for Data Science

For general data analysis and statistics guidance: * SPSS Tutorial is useful for general statistics knowledge * Stat Trek as a refresher of statistics concepts * Kent State SPSS Tutorials for seeing how data are traditionally processed * Kaggle Courses to see how machine learning can be used

References

[1] Biozentrum University of Basel, Switzerland

License

Distributed under the MIT License. See `LICENSE` for more information.