

1. Software for Modeling

Owen Hale

Model classes:

- Descriptive models: show trends in data
- Inferential models: test observed data against null hypothesis
- Predictive models: predict outcomes based on new data

Modeling terminology:

- **outcome**: dependent variable
- **predictor**: independent variable
- **supervised/unsupervised**: supervised models incorporate outcomes
 - **regression** is a category of supervised model that gives a numeric outcome
 - **classification** is another supervised model type that gives a categorical outcome

Modeling process:

- **Exploratory data analysis (EDA)**
- **Feature engineering**: applying EDA insights to producing a model
- **Model tuning and selection**: generate, optimize, and test different types of models
- **Model evaluation**: EDA for models