

---

# Stan for the people

## Two days introductory workshop on Bayesian modeling

---

**Instructor:** Charles Margossian  
charles.margossian@columbia.edu

**Teacher Assistant:** Aaron Erlich  
aaron.erlich@mcgill.ca

The course is divided into seven modular parts. The goal is to go through as many of them as possible, but our priority is to have an in depth discussion and spend time doing hands on exercises. You are given a course package, which includes the course slides, additional documentation, as well as R scripts and Stan files to do the exercises.

### Tentative outline

#### Day 1 - Morning

1. Introduction to Bayesian analysis
  - What is a model?
  - Bayesian inference
  - Beyond the posterior distribution
2. Algorithms and computational consideration
  - Monte Carlo methods
  - The geometry of posterior distributions
  - The Hamiltonian Monte Carlo sampler

#### Day 1 - Afternoon

3. Introduction to Stan
  - How Stan works
  - Inference algorithms
  - Example 1: linear regression
    - Live demo: writing a model in Stan
    - Diagnosing the inference algorithm
    - Posterior predictive checks
    - Live demo: the generated quantities block
    - *Exercise 1: Building an improved model*
  - Remarks on the prior distribution
  - General resources to use Stan
  - Example 2: logistic regression model
    - Review of the logistic model
    - *Exercise 2: build, fit, and diagnose a logistic model*
    - Posterior predictive checks for binary data
    - *Exercise 3: build, fit, and criticize a more complex logistic model*

## Day 2 - morning

4. Conversational Stan
  - Computing Markov chains in parallel
  - Variable types
  - Additional language blocks
5. Hierarchical models
  - Complete, partial, and no pooling
  - Principles behind a hierarchical model
  - Example 3: two way normal model
    - Proper and improper priors
    - *Exercise 3: write, fit, and diagnose a hierarchical model*

## Day 2 - Afternoon

6. Model parametrization
  - Example 4: Modeling Baseball players' hitting success
    - *Exercise 4: Analyze the estimates from the complete and no pooling models*
    - *Exercise 5: Write, fit, and analyze a partial pooling model*
    - Divergent transitions... and where they come from
    - *Exercise 6: Reparametrize the partial pooling model*
7. Open discussion and concluding remarks
  - What we covered
  - What we didn't covered
  - Where to learn more
  - The Stan software ecosystem
  - Other probabilistic programing languages