



PyMC is a probabilistic programming package for building Bayesian statistical models in Python. It provides flexible and performant implementations of cutting-edge model fitting algorithms, including gradient-based Markov chain Monte Carlo (MCMC) and variational inference (VI).

PyMC's intuitive model specification syntax makes it easy to build a wide range of models for any number of applications, making these powerful statistical tools accessible to those who may not be expert computer scientists. Moreover, its extensibility has allowed for a number of advanced extensions to be developed, which currently includes time series modeling, Gaussian processes, sequential Monte Carlo, and Bayesian additive regression trees (BART).

## APPLICATIONS

At HelloFresh, PyMC is used to optimize their marketing budget via Bayesian media mix modeling (MMM) and to conduct AB testing, which PyMC has helped speed up by more than 60x.

At U.K. broadcaster Channel 4, PyMC is used for developing internal forecasting tools.

PyMC is used as a tool for statistical modeling by a range of companies, including Roche, P&G, The Gates Foundation, Salesforce, Gain Theory, Monetate, Civiq, HelloFresh, and GrubHub.

## PLANNED FEATURES

- + **Upgrading Computational Backend:** The next version of PyMC will use *Aesara* as its backend, replacing Theano. This will result in large speed-ups due to symbolic optimizations of the computational graph. Switching to *Aesara* will also make it possible to use different libraries to provide the computation for model fitting, including JAX, NumPyro and Numba. With this, it becomes possible to leverage GPUs for improving the performance of PyMC's fitting algorithms.
- + **Modernizing Variational Inference:** As models and datasets get larger, variational inference becomes a more important tool for fitting Bayesian models. Since VI was originally implemented in PyMC, a number of methodological advances have been made which we would like to make available to users. We also want to be able to automate the evaluation of VI model outputs.
- + **Improved Gaussian Processes:** GPs are an invaluable tool for non-parametric modeling that makes Bayesian inference even more flexible. However, they tend to scale poorly with larger datasets; recent advances, such as black box matrix-matrix (BBMM) GPs allow for much better scaling, particularly when used in conjunction with GPUs.

# PROJECT NEEDS

Developer Support

\$10,000 - \$40,000/year

The PyMC project (pymc-devs on GitHub) now comprises 15-20+ active developers at any given time, contributing code, documentation, testing, and expertise to developing PyMC. Monetary support for our most productive developers would protect their time with the project and allow us to achieve our development goals in a timely fashion.

Community Support Manager

\$15,000/year

Our Discourse channel (discourse.pymc.io) is the primary avenue for support for most users. On Discourse, they can interact with both developers and other users to resolve usage problems or get questions answered without overburdening the project Issue Tracker on GitHub. However, the volume of traffic has gotten large enough to warrant a dedicated manager to ensure that users' needs are addressed reliably.



PyMC

For more information on PyMC, including our governance structure and project roadmap, please visit:

<https://docs.pymc.io/en/v3/>

PyMC is a Sponsored Project of NumFOCUS, a US 501(c)(3) public charity.

NumFOCUS Sponsored Projects rely on the generous support of corporate sponsors, institutional partners, and individual donors.

**NUMFOCUS**  
OPEN CODE = BETTER SCIENCE

For more information:

[info@numfocus.org](mailto:info@numfocus.org) | +1 (512) 831-2870