# Internship: Bayesian hierarchical modeling in marketing and sales

## Description

Marketing spends in Pharmaceuticals can have a considerable share in the overall product costs. Therefore, there is the need to identify the key drivers of the revenues from a mix of several marketing activities through so-called Marketing-Mix (MMx) models. Especially launch brands, however, operate in a low-data regime that hampers identifiability as opposed to established brands with a long history and therefore, strong data support.

In this internship, we will use adaptive Hamiltonian Monte-Carlo (HMC) sampling to estimate not only the effectiveness of promotional advertising but also the associated Bayesian credible intervals. We will work with algebraic, nonlinear, spatio-temporal models, and actual industrial data reflecting regionalized sales structures and networks between customers and sales force.

To inform launch through established brands, we will explore at least two options, a) multi-level hierarchical models, and b) the use of marginal posteriors of former campaigns as priors for future ones. The goal is a systematic comparison to identify the most expressive model.

To achieve this goal, we propose the following steps as part of this internship

* Implementation of a pool of models in Stan
* Literature research on dedicated model scores for hierarchical models to quantify both the gain in hierarchical modeling and model quality, and their implementation in R
* Prior predictive checks
* Tailoring (hyper)priors to control the trade-off between sampler convergence and model expressiveness

## Literature

[1] <https://storage.googleapis.com/gweb-research2023-media/pubtools/3804.pdf>

[2] <https://arxiv.org/abs/2008.12802>

## Eligibility

* Enrollment at TU Dortmund University required
* Internship for Module “Project Work”; this part of the module should not have been completed yet

## Must-haves

* Strong programming skills in R / RStudio, incl. data frames, data tables, dplyr, tidyr, ggplot
* Background in systematic model selection and maximum likelihood estimation
* Interest in Markov-Chain Monte-Carlo methods and Stan
* Interest in hierarchical (mixed effect) modeling
* Team mindset and creative modeling mindset

## Nice-to-haves

* Programming skills in python, incl. pandas, numpy, pyro, numpyro
* Knowledge of Stan syntax, rstan, HMC / NUTS, Gelman’s convergence criteria
* Background in multi-level hierarchical modeling
* Experience with Git, Linux

## Contacts

TU Dortmund University: Prof. Dr. Katja Ickstadt, Department of Statistics, ickstadt@statistik.tu-dortmund.de

Bayer AG: Dr. Moll Glass, Applied Mathematics, moll.glass@bayer.com

Applicants are asked to include their CV, preferred start and end date and preferred workload.

## Miscellaneous

We recommend a duration of at least 3 months at 80% (4 days internship, 1 day off). This is a paid internship. Bayer will also provide the necessary hardware and a shared office. Working from home is possible.