# Marta Karas


CONTACT INFORMATION

929 N Wolfe St Apt 607D
Baltimore, MD 21205
USA

*Voice:* +1 317-665-4551
*E-mail:* marta.karass@gmail.com
*Web:* https://martakarass.github.io/


EDUCATION

**Johns Hopkins Bloomberg School of Public Health**, Baltimore, MD, USA

***Ph.D., Biostatistics***, Aug 2017–present

- Academic Advisors: Ciprian Craniceanu, Jacek K. Urbanek

**Wroclaw University of Science and Technology**, Wroclaw, Poland

***M.S., Mathematics*** (Mathematical Statistics), Jul 2015

- Dissertation: "Theoretical and practical issues in change point detection."
- Academic Advisor: Malgorzata Bogdan
- Final grade: 5.5 (Excellent). Graduation with Academic Distinction

**Wroclaw University of Science and Technology**, Wroclaw, Poland

***B.S., Mathematics***, Jun 2013

- Final grade: 5.0 (Very good)

HONORS AND AWARDS

1. ENAR Poster Award. 2017

ACADEMIC EXPERIENCE

**Indiana University Bloomington**, Department of Epidemiology and Biostatistics, School of Public Health, Bloomington, IN, USA

*Research Assistant* **Jan 2017 - Jul 2017**

- Applying graph-constrained regularization methods to determine what brain structural imaging markers are associated with HIV+/HIV− status.

**Indiana University – Purdue University Indianapolis**, Department of Biostatistics, Richard M. Fairbanks School of Public Health, Indianapolis, IN, USA

*Research Assistant* **Jan 2016 - Jul 2016**

- Developing extension of existing graph-constrained regularization methods for linear regression. Determining what brain structural imaging markers are associated with alcohol abuse.

PUBLICATIONS

1. **Karas, M.**, Straczkiewicz, M., Fadel, W., Harezlak, J., Crainiceanu, C., Urbanek, J.K. *Adaptive empirical pattern transformation (ADEPT) with application to walking stride segmentation*, Submitted to *Biostatistics*, 2018.
2. **Karas, M.**, Bai, J., Straczkiewicz, M., Harezlak, J., Glynn, N W., Harris, T., Zipunnikov, V., Crainiceanu, C., Urbanek, J.K. *Accelerometry data in health research: challenges and opportunities. Review and examples*, Submitted to Statistics in Biosciences, 2017. (*preprint*)
3. Brzyski, D., **Karas, M.**, Ances, B., Dzemidzic, M., Goni, J., Randolph, T.W., Harezlak, J. *Connectivity-Informed Adaptive Regularization for Generalized Outcomes*, 2017. (*preprint*)
4. **Karas, M.**, Brzyski, D., Dzemidzic, M., Goni, J., Kareken, D.A., Randolph, T., Harezlak, J. *Brain connectivity-informed regularization methods for regression*, Statistics in Biosciences, 2017. (*link*)

## Conference Posters

1. **Karas, M.**, Brzyski, D., Ances, B., Goni, J., Randolph, T.W., Harezlak, J. *Association of Structural Brain Imaging Measures with HIV Markers Incorporating Structural Connectivity Information: a Regularized Statistical Approach.* ENAR, Washington DC, USA, Mar 2017. (Received ENAR Poster Award).
2. **Karas, M.** *Penalized regression inference regarding variable selection in high dimensions: presentation of selected methods implemented in R.* European R Users Conference, Poznan, Poland, Oct 2016.

## Oral Presentations

1. **Karas, M.**, Harezlak, J., Straczkiewicz, M., Fadel, W., Crainiceanu, C., Urbanek, J.K. *ADaptive Empirical Pattern Transformation (ADEPT) with application to walking stride segmentation.* JSM 2018, Vancouver, Canada, Aug 2018.
2. **Karas, M.**. *Wearable accelerometers, accelerometry data and automatic steps segmentation in R: strideter and convo R packages.* Why R? 2018 Conference, Wroclaw, Poland, Jul 2018.
3. **Karas, M.** *Bayesian analysis with R and Stan: introduction and application to a business case problem.* Wroclaw Data Science Meetup, Wroclaw, Poland, Nov 2016.
4. **Karas, M.**. *Convex Clustering and Biclustering with application in R.* Cracow R Users Meetup, Krakow, Poland, Sep 2016.

## Software

`mdpeer` R package: Graph-Constrained Regression with Enhanced Regularization Parameters Selection. (*link*)

## Industry Experience

**Opera Software** (Software), Wroclaw, Poland

*Analyst* **Aug 2016 - Dec 2016**

- Developed time series forecasting model for Opera browser's core metrics.
- Performed Bayesian analysis to describe and infer about browser's users performance.

**Opera Software** (Software), Wroclaw, Poland

*Analyst* **Jul 2015 - Dec 2015**

- Established methodology for A/B-test results analysis, including selection of statistical methods and implementation of a complete tool in `R`.
- Performed user base analysis for software product improvements.

**Datarino** (Big Data Services & Data Management Solutions), Wroclaw, Poland

*Data Scientist* **Jul 2014 - Mar 2015**

- Analyzed user activity and monetization KPIs of a Polish social networking service.
- Retrieved knowledge from business partners' big-data size data sets.

**KRUK S.A.** (Debt collection), Wroclaw, Poland

*Intern* **Jun 2014**

- Compared feature selection methods addressing serious correlation problem.
- Implemented and applied a text mining tool for utilizing unstructured text data in `R`.

**QuantUp** (Data analysis, modeling and training), Wroclaw, Poland

*Intern* **Jul 2013 - Apr 2014**

- Performed research and case-study analysis of building and validating scoring models.
- Wrote articles about reproducible research tools and large-size data analysis tools in `R`.

## Computer Skills

- Languages / Tools:
  - Experienced: `R`, `Python`, `SQL`.
  - Familiar: `Matlab`.
- Operating Systems: Linux, OS X, Windows.

## Competitions and Hackatons

1. 2nd place in Data Analysis Marathon: "Determining tabloidization index". Krakow, Poland, Nov 2015.
2. Participation in "HackZurich // The biggest European hackathon". Zurich, Switzerland, Oct 2015.
3. The Winner of the Schneider Electric Business Analyst Competition 2015. Wroclaw, Poland, Jun 2015.
4. 2nd place in the Independent National Data Analysis Competition "NOMAD" V edition 2015. Wroclaw, Poland, Jun 2015.
5. 2nd place in the Independent National Data Analysis Competition "NOMAD" IV edition 2014. Wroclaw, Poland, Jun 2014.
6. The Winner of the KRUK Analytics Challenge 2014. Wroclaw, Poland, May 2014.

## Independent Coursework

- University of Washington: *Summer Institute in Statistics for Big Data (SISBID)*, Jul 11 - 27, 2016, Seattle, WA, USA:
  1. Reproducible Research for Biomedical Big Data.
  2. Supervised Methods for Statistical Machine Learning.
  3. Unsupervised Methods for Statistical Machine Learning.
- Johns Hopkins University & University of Colorado Boulder: *Principles of fMRI 1* (online via Coursera).
- Stanford University: *Mining Massive Datasets* (online via Coursera).
- MIT: *Introduction to Computer Science and Programming Using Python* (online via edX).