

BRUNA WUNDERVALD

@brunadaviesw@gmail.com
@bwundervald

+353 087 1877 482
github.com/brunaw

Hamilton Institute, Maynooth, Ireland

http://brunaw.com/

SUMMARY

I am a Statistician, researcher, and an open-source programming languages enthusiast. I like statistical modelling and machine learning in general, package and dashboard development, and teaching. I'm specially interested in Bayesian inference, multivariate analysis, feature extraction, probabilistic graphical models and natural language processing.

Currently, I am pursuing my Ph.D. at the Hamilton Institute in Maynooth, Ireland. I'm part of Andrew Parnell's research group, and our work is mostly related to the development of new Bayesian methods for machine learning and its many possible applications.

WORK EXPERIENCE

Statistician and Data Analyst, **Atturos**

2020 Dublin, Ireland

- Data analysis & machine learning for clinical experiments.

Research Assistant, **Hamilton Institute**

2020 Maynooth, Ireland

- Data analysis about COVID-19 in Ireland.

Graduate Teaching Assistant, **Maynooth University**

2019 Maynooth, Ireland

- Lab conducting, student support and paper grading for MSc students.

Lead Statistician, **Priced**

2018 São Paulo, São Paulo, Brazil

- Analysis of expenditure data in Brazil: data wrangling, product elasticity and churn, consumers clusterization and general behaviour.

Statistics Intern, **City Hall of Pinhais**

2017 Pinhais, Paraná, Brazil

- Geostatistical data analysis on the financial data of the people living in Pinhais City.

Statistics Intern, **Bradesco Bank**

2016-2017 Curitiba, Paraná, Brazil

- Credit risk modelling with logistic regression.

STRENGTHS

Bayesian Methods Multivariate Statistics Variational Inference

MIR Machine Learning NLP

R tidyverse python julia c++ Shiny JAGS & Stan

git GitHub HTML CSS L^AT_EX

EDUCATION

Ph.D. candidate in Statistics **Maynooth University**

Sept 2018 - *

Thesis theme: Extensions to Bayesian Tree-Based Models

Honours Bachelor degree in Statistics **Federal University of Paraná**

April 2013 - August 2018

Thesis theme: Machine learning and chord based feature engineering for genre prediction in popular Brazilian music

PROJECTS

R-Ladies Organizer
São Paulo, Curitiba and Dublin,
<https://rladies.org/>

R-Music Blog
<https://r-music.rbind.io/>

Non-Linear Models Catalog
<https://github.com/leg-ufpr/nlmSet>

LANGUAGES

English ●●●●●

Portuguese ●●●●●

French ●●●●●

Spanish ●●●●●

Italian ●●●●●

Dutch ●●●●●

Russian ●●●●●

REVIEWER

- **rOpenSci** - R Packages Reviewer. 2018 - Current.
- **TEMA: Trends in Applied and Computational Mathematics** - Journal Articles Reviewer. 2019 - Current.
- **ISMIR 2020** - Conference Reviewer. 2020 - Current.
- **JOSS** - Journal Articles Reviewer. 2020 - Current.

PUBLICATIONS

Books

- PET Estatística UFPR. *GIT Manual, Project of Tutorial Extension in Statistics*. 2016.

Articles

- B. Wundervald, A. C. Parnell, and K. Domijan. "Generalizing Gain Penalization for Feature Selection in Tree-Based Models". In: *IEEE Access* 8 (2020), pp. 190231–190239. DOI: [10.1109/ACCESS.2020.3032095](https://doi.org/10.1109/ACCESS.2020.3032095).
- Bruna Wundervald. "Cluster-based quotas for fairness improvements in music recommendation systems". In: *International Journal of Multimedia Information Retrieval* (2021), pp. 1–8.
- Bruna Wundervald. "Feature Engineering for Genre Characterization in Brazilian Music". In: *Conference: 13th International Workshop on Machine Learning and Music* (Sept. 2020), pp. 60–65. URL: <https://rafaelramirez8.wixsite.com/mml2020>.
- Bruna D. Wundervald and Walmes M. Zeviani. "Machine learning and chord based feature engineering for genre prediction in popular Brazilian music". In: (Feb. 2019). arXiv: [1902.03283](https://arxiv.org/abs/1902.03283). URL: <http://arxiv.org/abs/1902.03283>.

Software

- **vagalumer: Access to the 'vagalume' api** - Author, [Online] Available: <https://CRAN.R-project.org/package=vagalumer>
- **chorrds: Music chords extraction** - Author, [Online] Available: <https://CRAN.R-project.org/package=chorrds>
- **broom: Convert Statistical Analysis Objects into Tidy Tibbles** - Contributor, [Online] Available: <https://github.com/tidymodels/broom>
- **distributions3: Probability Distributions as S3 Objects** - Contributor, [Online] Available: <https://github.com/alexpgayes/distributions3>
- **ranger: A Fast Implementation of Random Forests** - Contributor, [Online] Available: <https://github.com/imbs-hl/ranger>

Miscellaneous

- B. Wundervald. *R-Music: Introduction to the vagalumeR package*. 2018. URL: <https://r-music.rbind.io/posts/2018-11-22-introduction-to-the-vagalumer-package/>.
- B. Wundervald and T. M. Dantas. *R-Music: Rspotify*. 2018. URL: <https://r-music.rbind.io/posts/2018-10-01-rspotify/>.
- B. Wundervald, F.A. Lima, and S.L.P. Ferrari. *Distributions on the Unit Interval Catalog*. Shiny Application. 2018. URL: <https://brunaw.shinyapps.io/dissertacao-francimario/>.
- B. Wundervald and J. Trecenti. *R-Music: Introduction to the chorrds package*. 2018. URL: <https://r-music.rbind.io/posts/2018-08-19-chords-analysis-with-the-chorrds-package/>.
- B. Wundervald and J. Trecenti. *R-Music: music21*. 2018. URL: <https://r-music.rbind.io/2018-10-06-music21>.

TEACHING EXPERIENCE

- | | |
|-----------|--|
| 2020/2021 | Statistical Methods for Data Science 1 & 2, Tutoring instructor, <i>Maynooth University</i> , 1st Term. |
| 2020 | Dados 360, NLP Course Instructor, <i>Escola de Dados</i> , online. |
| 2019/2020 | Topics in Data Analytics, Tutoring instructor, <i>Maynooth University</i> , 2nd Term. |
| 2019/2020 | R for Data Analytics, Tutoring instructor, <i>Maynooth University</i> , 1st Term. |
| 2019/2020 | Statistical Methods for Data Science, Tutoring instructor, <i>Maynooth University</i> , 1st Term. |
| 2019 | Music Data Analysis in R, Shortcourse instructor, <i>IV International Seminar on Statistics with R</i> , May 2019. |
| 2019 | Building R packages Workshop, Workshop teaching assistant, <i>University College Dublin</i> , April 2019. |
| 2019 | Deep Learning for Statisticians, Shortcourse instructor, <i>XVI School of Regression Models</i> , March 2019. |
| 2019 | Statistical Machine Learning, Tutoring instructor, <i>Maynooth University</i> , 1st Term. |
| 2018 | Linear Models 2, Tutoring instructor, <i>Maynooth University</i> , 2nd Term. |
| 2017 | R Markdown, Shortcourse instructor, <i>Federal University of Paraná</i> September 2017. |

- B. Wundervald and W.M. Zeviani. *Non-Linear Models Catalog*. Shiny Application. 2017. URL: <https://github.com/brunaw/catreg>.
- Bruna Wundervald. *R-Music: Connecting packages*. 2019. URL: <https://r-music.rbind.io/posts/2019-03-13-connecting-packages/>.

CONFERENCE & INVITED TALKS

- 2020 Mixture Cure Rate Models for the Analysis of Survival Times in the COVID-19 Scenario
Hamilton Institute Students Seminar Series,
Online, November 2020.
- 2020 How can we forecast epidemic diseases?
Maynooth University Research Week,
Online, October 2020.
- 2020 Feature Engineering for Genre Characterization in Brazilian Music
13th International Workshop on Machine Learning and Music,
Online, September 2020.
- 2019 The tidyverse for Machine Learning
satRday São Paulo,
São Paulo, Brazil, November 2019.
- 2019 Probabilistic Graphical Models in R and python
IV International Seminar on Statistics with R,
Niterói, Brazil, May 2019.
- 2019 Regularization Methods in Random Forests
39th Conference on Applied Statistics in Ireland,
Dundalk, Ireland, May 2019.
- 2019 Chord Based Feature Extraction for Genre Classification in Popular Brazilian Music,
XVI School of Regression Models,
Pirenópolis, Brazil, March 2019.
- 2018 Construction and Implementation of Multivariate Dispersion Models,
63rd meeting of the Brazilian Region of the IBS
Curitiba, Brazil, May 2018.
- 2018 Music Data Analysis with R,
R Day - First national meeting of R users
Federal University of Paraná
Curitiba, Brazil, May 2018.
- 2016 Non-Linear Models Catalog,
Integrated Week of Teaching, Research and Extension,
Federal University of Paraná
Curitiba, Brazil, 2016.
- 2015 Results of the Applied Statistical Analysis Laboratory,
Integrated Week of Teaching, Research and Extension,
Federal University of Paraná,
Curitiba, Brazil, 2015.

ACADEMIC WORK

- 2021 Modeling Session, co-host,
rstudio::conf, *Online*
- 2019/2020 Students Seminar Series,
Chair & Organizer,
Hamilton Institute,
Maynooth University
- 2019 *satRdays* São Paulo,
Scientific Committee & Organizer,
São Paulo, São Paulo, Brazil.
- 2019 *PyCon Ireland 2019*,
Session Chair,
Dublin, Ireland.

POSTER PRESENTATIONS

- 2019 Regularization Methods in Random Forests,
Machine Learning Summer School,
London, England, July 2019.
- 2019 Construction and Implementation of Multivariate Dispersion Models,
XVI School of Regression Models,
Pirenópolis, Brazil, March 2019.
- 2018 Chord Based Feature Extraction for Genre Classification in Popular Brazilian Music, *1st WiMIR*,
Paris, France, September 2018.
- 2017 Non-Linear Models Catalog,
XV School of Regression Models,
Goiânia, Brazil, March 2017.
- 2016 The use of R at Federal University of Paraná, *I International Seminar of Statistics with R*,
Niterói,
Rio de Janeiro, Brazil, May 2016.

SEMINARS AND MEETUPS

- 2021 An introduction to the `tidymodels` package, *Young ISA*, Online, January 2021.
- 2020 The `tidyverse` for Machine Learning, *R-Ladies Helsinki Meetup*, Online, June 2020.
- 2020 Introduction to `Julia` Programming, *Hamilton Group*, Online, May 2020.
- 2020 I'm a Data Scientist, Git me out of here!, *R-Ladies Dublin Meetup*, Dublin, Ireland, February 2020.
- 2019 Tree based models in `R`, *R-Ladies Dublin Meetup*, Dublin, Ireland, June 2019.
- 2019 An Introduction to Bayesian Regression Trees *Federal University of Paraná*, Curitiba, Brazil, March 2019.
- 2019 `googleAnalyticsR` & `purrr`, *R-Ladies São Paulo Meetup*, São Paulo, Brazil, March 2019.
- 2018 The steps of a Kaggle project, *Hamilton Institute* Maynooth, Ireland, November 2018.
- 2018 Git & GitHub tutorial, *Hamilton Institute* Maynooth, Ireland, October 2018.
- 2018 Basic `R` + `dplyr`, *R-Ladies São Paulo Meetup*, São Paulo, Brazil, August 2018.
- 2018 Basic `R` + `dplyr`, *R-Ladies Curitiba Meetup*, Curitiba, Brazil, June 2018.

TRAVEL & DIVERSITY GRANTS

- 2020 *ISMIR 2020*, Conference Scholarship, Online.
- 2019 *ISMIR 2019*, Conference Scholarship, Delft, The Netherlands.
- 2019 *PyCon Ireland 2019*, Conference Scholarship, Dublin, Ireland.
- 2019 *StanCon 2019*, Conference Scholarship, Cambridge, England.
- 2019 *Women in Data Science*, Travel & Conference Scholarship, Zurich, Switzerland.
- 2019 *Rstudio Conference*, Travel & Conference Scholarship, Austin, Texas.

GRANTS

- 2017 Scientific Initiation, supervised by Prof. PhD Wagner Hugo Bonat, *Construction of Multivariate Probability Distributions*. Duration: 12 months
- 2016 Project of Tutorial Extension in Statistics, supervised by Prof. Dr. Walmes Marques Zeviani, *Non-Linear Models Catalog*. Duration: 10 months
- 2015 Laboratory of Applied Statistics, supervised by Prof. Dr. Cesar Augusto Taconeli and Prof. Dr. Fernanda Buhner Rizzato. Duration: 6 months
- 2014 Teaching Initiation Program, supervised by Prof. Dr. Jomar Camarinha Filho Course: Basic Statistics. Duration: 4 months
- 2014 Teaching Initiation Program, supervised by Prof. Dr. Jomar Camarinha Filho Course: Statistics II. Duration: 4 months

SHORT COURSES ATTENDED

- 2019 Design of Experiments and Studies, *Oxford University*, 12h
- 2019 Flexible Regression, *Oxford University*, 12h
- 2019 Model assessment and selection, *Cambridge University*, 3h
- 2019 A Dive into Stan's C++ Model Concept, *Cambridge University*, 3h
- 2019 Hierarchical Modeling with `Stan`, *Cambridge University*, 6h
- 2019 Machine Learning Summer School, *University College London*, 40h
- 2019 High-dimensional Statistics, *Durham University*, 12h
- 2019 Computer Intensive Statistics, *Durham University*, 12h
- 2019 Shiny, *IV International Seminar on Statistics with R*, 3h
- 2019 Statistics in the Cloud, *Dundalk*, 3h
- 2019 Statistical Modelling, *University of Southampton*, 12h
- 2019 Applied Stochastic Processes, *University of Southampton*, 12h
- 2019 Applied Machine Learning, *Rstudio Conference 2019*, Austin, Texas, 12h
- 2018 Statistical Computing, *Cambridge University*, 12h
- 2018 Statistical Inference, *Cambridge University*, 12h
- 2018 Web Scraping with R, *Curso-R*, 8h.
- 2018 Introduction to Cognitive Diagnostic Models with Stan, *Federal University of Paraná*, 2h.
- 2018 Distributed data structures with R, *Federal University of Paraná*, 2h.
- 2018 Dashboards with R, *Curso-R*, 4h.
- 2018 Probability Theory, *University of São Paulo*, 60h.
- 2017 Machine Learning with R, *University of São Paulo*, 4h.
- 2017 APIs with R, *University of São Paulo*, 4h.
- 2017 Multivariate Covariance Generalized Linear Models for the Analysis of Experimental Data, *Federal University of Lavras*, 4h.
- 2017 Regression models for count data: beyond the Poisson model, *Goiânia, GO, Brazil*, 4h.
- 2017 Elements of Computational Statistics, *Federal University of São Carlos*, 6h.
- 2016 Modelling with SAS, *HSBC Bank Brazil*, 20h
- 2016 Statistical Modelling Academy, *HSBC Bank Brazil*, 192h.
- 2015 Conversation in English, *English without Borders*, *Federal University of Paraná*, 16h.
- 2015 English for Professional and Academic Goals, *English without Borders*, *Federal University of Paraná*, 16h.
- 2015 English with Music, *English without Borders*, *Federal University of Paraná*, 16h.
- 2015 Advanced Excel, *Center of School-Enterprise Integration of Paraná*, 20h.

REFEREES

Prof. PhD Andrew Parnell

@ Hamilton Institute
✉ andrew.parnell@mu.ie
Maynooth University,
+353 1 716 2422,

Prof. PhD Wagner Hugo Bonat

@ Department of Statistics
✉ wbonat@ufpr.br
Federal University of Paraná,
+55 41 3361-3254,

Prof. PhD Paulo Justiniano Ribeiro Júnior

@ Department of Statistics
✉ paulojus@ufpr.br
Federal University of Paraná,
+55 41 3361-3254,

Prof. Dr. Victor Fossaluzza

@ Institute of Mathematics and Statistics
✉ victorf@ime.usp.br
University of São Paulo
+55 11 3091-6187