Guillermo Vinué, PhD, Data scientist

Personal and contact information

25/10/1985, Valencia, Spain
Phone number: +34 659 363 291
ID Card: 22 58 95 73 P

E-mail: guillermovinue@gmail.com
Website: `https://www.uv.es/vivigui`
GitHub: `https://github.com/guivivi`

Education

PhD in Statistics and Optimization
University of Valencia, Spain

September 2010 – September 2014

Master's degree in Biostatistics
University of Valencia, Spain

October 2010 – June 2012

Degree in Mathematics
University of Valencia, Spain

September 2003 – June 2008

Awards and fellowships

PhD extraordinary award

Master's degree extraordinary award

Selected to represent Spain in the 18th European Young Statisticians Meeting

2nd prize in the *Early-Stage Researcher Prize* of the XIV Biometric Spanish Conference

Erasmus Scholarship
Johannes Gutenberg University, Mainz, Germany

September 2007 – June 2008

Professional experience

Company related to military econometrics
Valencia, Spain (telework)
Data analyst, Data scientist, Statistician
Details: Statistical support and R programming with econometrics data.

February 2024 – Nowadays

Akkodis Technologies Spain S.L.U.
Valencia, Spain (telework)
Data analyst, Data scientist, Statistician
Details: Statistical support and R programming for GSK (biopharma company).

August 2022 – June 2023

Symanto Spain S.L.
Valencia, Spain
Data analyst, Data scientist, Statistician
Details: Statistical support, natural language processing and R programming for Human Resources companies.

February 2020 – March 2022

Tyris Software S.L.
Valencia, Spain   April 2019 – December 2019
Data analyst, Data scientist, Statistician
Details: Statistical support and R programming for Sacyr and Acciona (infrastructures companies).

Department of Computer Science
KU Leuven, Leuven, Belgium   March 2018 – March 2019
Data analyst, Data scientist, Statistician, Postdoctoral researcher
Details: Anomaly detection on time-series data and R programming.

Vienna Institute of Demography (VID)
Austrian Academy of Sciences, Vienna, Austria   March 2016 – March 2018
Data analyst, Data scientist, Statistician, Postdoctoral researcher
Details: Statistical modeling of human migration data and R programming.

UCD School of Mathematics and Statistics/The Insight Centre for Data Analytics
University College Dublin (UCD), Dublin, Ireland   February 2015 – February 2016
Data analyst, Data scientist, Statistician, Postdoctoral researcher
Details: Statistical classification of online products and R programming.

Department of Statistics and Operational Research
University of Valencia, Valencia, Spain   September 2010 – September 2014
Data analyst, Data scientist, Statistician, PhD student
Details: Completion of the doctoral thesis and R programming.

Drug Research Center
Santa Creu i Sant Pau Hospital, Barcelona, Spain   March 2009 – March 2010
Data analyst, Data scientist, Statistician, Research technician
Details: Data management and R programming.

Language skills

- **Spanish**: Mother tongue

- **English**: C1, certified by Cambridge Assessment English

- **German**: C2 (Grosses Deutsches Sprachdiplom), certified by Goethe Institut
  Other certificates: C1 (Zentrale Mittelstüfen Prüfung); B1 (Zertifikat Deutsch), both certified by Goethe Institut

- **French**: B2, certified by DELF from the French Ministry of Education
  Other certificates: *Consolidated basic knowledge*, certified by from Alison MOOC; *Parler le français des affaires*, certified by GoodHabitz online training

---

Technical skills

- **Advanced R**: data processing, analytics and visualisation programming (tidyverse, shiny, flexdashboard, tidymodels, roxygen2, knitr, ...)

- **Other Languages, Database Tools and Operating Systems**: LaTeX, HTML, Python, SQL, Matlab, SPSS, C++, Linux, Windows, MS Office

Publications (Papers and software)

- Vinué G., "A Basketball Big Data Platform for Box Score and Play-by-Play Data", *Big Data 3*, 2024, **Impact Factor: 4.6, JCR 2022),** https://doi.org/10.1089/big.2023.0177

- Vinué G. et al.**(+1)**, "Robust archetypoids for anomaly detection in big functional data", *Advances in Data Analysis and Classification 15(2), 437-462*, 2021, **Impact Factor: 2.134, Position: 40/125 = 0.32 (Statistics & Probability, JCR 2020), Cites: 19**, https://doi.org/10.1007/s11634-020-00412-9

- Vinué G., "A web application for interactive visualization of European basketball data", *Big Data 8(1), 70-86*, 2020, **Impact Factor: 2.128, Position: 37/110 = 0.34 (Computer Science, Theory & Methods, JCR 2020), Cites: 7**, https://doi.org/10.1089/big.2018.0124

- Vinué G. et al.**(+1)**, " Forecasting basketball players' performance using sparse functional data", *Statistical Analysis and Data Mining: The ASA Data Sci Journal 12(6), 1-14*, 2019, **Impact Factor: 1.396, Position: 50/124 = 0.4 (Statistics & Probability, JCR 2019), Cites: 20**, https://doi.org/10.1002/sam.11436

- Vinué G. et al.**(+1)**, "Archetypoid Analysis for Sports Analytics", *Data Mining and Knowledge Discovery 31(6), Special Issue on Sports Analytics, 1643-1677*, 2017, **Impact Factor: 2.481, Position: 40/132 = 0.3 (Computer Science, Artificial Intelligence, JCR 2017), Cites: 32**, https://doi.org/10.1007/s10618-017-0514-1

- Vinué G., "Anthropometry: An R Package for Analysis of Anthropometric Data", *Journal of Statistical Software 77(6), 1-39*, 2017, **Impact Factor: 22.737, Position: 1/123 (Statistics & Probability, JCR 2017), Cites: 32**, https://doi.org/10.18637/jss.v077.i06

- Vinué G. et al.**(+2)**, "The k-means algorithm for 3D shapes with an application to apparel design", *Advances in Data Analysis and Classification 10(1), 103-132*, 2016, **Impact Factor: 2.326, Position: 11/124 = 0.09 (Statistics & Probability, JCR 2016), Cites: 24**, https://doi.org/10.1007/s11634-014-0187-1

- Vinué G. et al.**(+2)**, "Archetypoids: a new approach to define representative archetypal data", *Computational Statistics and Data Analysis 87, 102-115*, 2015, **Impact Factor: 1.179, Position: 49/123 = 0.4 (Statistics & Probability, JCR 2015), Cites: 37**, https://doi.org/10.1016/j.csda.2015.01.018

- Vinué G. et al.**(+3)**, "Looking for representative fit models for apparel sizing", *Decision Support Systems 57, 22-33*, 2014, **Impact Factor: 2.313, Position: 11/81 = 0.14 (Operations Research & Management Science, JCR 2014), Cites: 30**, https://doi.org/10.1016/j.dss.2013.07.007

- Epifanio I., Vinué G. et al.**(+1)**, "Archetypal analysis: Contributions for estimating boundary cases in multivariate accommodation problem", *Computers & Industrial Engineering 64, 757-765*, 2013, **Impact Factor: 1.783, Position: 11/43 = 0.26 (Engineering, Industrial, JCR 2013), Cites: 31**, https://doi.org/10.1016/j.cie.2012.12.011

- Ibáñez M. V., Vinué G. et al.**(+5)**, "Apparel sizing using trimmed PAM and OWA operators", *Expert Systems with Applications 39, 10512-10520*, 2012, **Impact Factor: 1.854, Position: 13/79 = 0.16 (Operations Research & Management Science, JCR 2012), Cites: 22**, https://doi.org/10.1016/j.eswa.2012.02.127

- Vinué G. et al.**(+1)**, "adamethods: Archetypoid Algorithms and Anomaly Detection. R package version 1.2.1", **R software repository CRAN**, 2020, https://CRAN.R-project.org/package=adamethods

- Vinué G., "BAwiR: Analysis of Basketball Data. R package version 1.3.2", **R software repository CRAN**, 2021, https://CRAN.R-project.org/package=BAwiR

- Vinué G. et al.**(+5)**, "Anthropometry: An R Package for Analysis of Anthropometric Data. R package version 1.19", **R software repository CRAN**, 2021, `https://CRAN.R-project.org/package=Anthropometry`

Software licences

- "User Interface (UI) for Prediction Models from IP arising from EI IP/2014/0345"
  (UCD Licence Agreement)

- "Machine Learning Algorithms (for Clavis) from IP arising from EI IP/2014/0345"
  (UCD Licence Agreement)

Revisions

- **Peer review:** Web of Science profile

- **Books:** (i) Book proposal for the "Elsevier's book series in textiles technology" ; (ii) Book proposal for the "John Wiley & Sons editorial"

- **Conferences:** Program committee member of MLSA17, MLSA18, MLSA19 and MLSA20 (The Machine Learning and Data Mining for Sports Analytics workshop)

Scientific conferences

- **UEG 2023**. 40th Summer School of the University of Valencia, Gandía, Spain, 13/07/2023.
  Invited speaker

- **SYSORM 2017**. 1st Spanish Young Statisticians and Operational Researchers Meeting, Granada, Spain, 13-15/11/2017.
  Invited speaker

- **BSPS 2016**. British Society for Population Studies, Winchester, England, 12-14/09/2016.
  Speaker

- **CASI 2015**. 35th Conference on Applied Statistics in Ireland, Cork, Ireland, 11-13/05/2015.
  Speaker

- **EYSM 2013**. 18th European Young Statisticians Meeting, Osijek, Croatia, 26-30/08/2013.
  Speaker. Recording of the presentation

- **CEB 2013**. XIV Biometric Spanish Conference, Ciudad Real, Spain, 22-24/05/2013.
  Speaker in the session of early-stage researchers

- **COMPSTAT 2012**. International Conference on Computational Statistics, Limassol, Cyprus, 27-31/08/2012. *Speaker*

Some dashboards

- Basketball dashboard

- Movies dashboard

Some outreach activities

- Producto nacional a la baja. Gigantes del Basket (in the press, number 1522)

- https://www.sportsandvisualdata.com

Teaching

Classes	Degree	Place	Duration	Hours
Introductory Statistics Using R for Computational Biologists	Bioinformatics and Systems Biology PhD Programme	UCD	11/2015 – 12/2015	30
Practical classes of Introduction to the Statistical Analysis and Data Processing	Information and Documentation	Faculty of Geography and History (UV)	02/2014 – 05/2014	30
Practical classes of Statistics	Medicine	Faculty of Medicine and Dentistry (UV)	09/2013 – 12/2013	27
Practical classes of Statistics	Medicine	Faculty of Medicine and Dentistry (UV)	09/2012 – 12/2012	31

Some courses

- Deep Learning With Python
  University of Sevilla. In progress

- The Complete Agile Scrum Fundamentals Course
  Udemy. 13 hours. August 2023

- Statistical Quality Control With R
  University of Vigo. 30 hours. February 2022

- Text Mining
  Madrid UPM Advanced Statistics and Data Mining Summer School. 15 hours. July 2015

- LFS101x: Introduction to Linux
  MOOC offered by The Linux Foundation through edX. 50 hours. August 2014

Other achievements

Accredited by the *National Quality Assessment and Accreditation Agency of Spain (ANECA)* as an *assistant lecturer*, *temporary lecturer* and *private university lecturer*

Spanish driving licence (B)

Participation in the 2007 Valencia Marathon (3h:53m)

References

- Irene Epifanio. PhD supervisor. University Jaume I of Castellón.
  epifanio@mat.uji.es

- Amelia Simó Vidal. PhD supervisor. University Jaume I of Castellón.
  simo@mat.uji.es