

# Efthymios Costa

Office 6M50, Huxley Building

Imperial College London, Department of Mathematics  
180 Queen's Gate, SW7 2AZ London, UK

(+44)7938778387

ec1917@ic.ac.uk

GitHub Profile

Personal Website

## EDUCATION

---

### PhD in Statistics and Machine Learning (StatML)

Oct 2021 - Sep 2025

Imperial College London, Department of Mathematics

Supervisors: Dr Ioanna Papatsouma & Professor Alastair Young

Thesis title: *Novel unsupervised techniques for mixed-type data*

### Visiting Doctoral Student

Jan - Mar 2024

University of Bologna, Department of Statistical Sciences "Paolo Fortunati"

Visiting Supervisor: Professor Christian Martin Hennig

### MSc Statistics (Data Science)

Sep 2020 - Sep 2021

Imperial College London, Department of Mathematics

Distinction (80.89/100)

Thesis: "A benchmarking study of distance-based clustering methods for mixed-type data"

### BSc Mathematics

Sep 2017 - Jun 2020

Imperial College London, Department of Mathematics

First Class Honours (78.21/100)

## EMPLOYMENT

---

### Research Associate with Teaching Duties in Statistics

Oct 2025 - Present

Imperial College London, Department of Mathematics

London, United Kingdom

Postdoc & Fellows Representative in the Department of Mathematics

### Research Intern

Jun - Aug 2020

The Cyprus Institute, Climate & Atmosphere Research Centre

Nicosia, Cyprus

Project title: "Climate Impacts on Vector-borne Diseases: Integrating Environmental Drivers to the Mathematical Models of Vector Populations"

### Advisory, Data Analytics Intern

Jun - Jul 2019

PwC Cyprus

Nicosia, Cyprus

## RESEARCH INTERESTS

---

Unsupervised learning: Cluster analysis, outlier detection, dimensionality reduction.

Mixed-type data: Combination of continuous and discrete variables in data.

Robust statistics: Robustness guarantees for machine learning methodologies.

## PUBLICATIONS & PREPRINTS

---

**Efthymios Costa**, Christian Martin Hennig (2026). "A unified approach to outlier identification for mixed-type data" (*manuscript in preparation*).

**Efthymios Costa**, Ioanna Papatsouma, Angelos Markos (2026). "Sparse clustering via the Deterministic Information Bottleneck algorithm". *arXiv preprint arXiv:2601.20628*.

**Efthymios Costa**, Ioanna Papatsouma (2025). "A novel framework for quantifying nominal outlyingness". In: *Statistics and Computing* 36, pp. 41–58.

**Efthymios Costa**, Ioanna Papatsouma (2025) Contribution to the discussion of Dümbgen and Davies (2025) "Connecting Model-Based and Model-Free Approaches to Linear Least Squares Regression". In: *Statistica* 84, pp. 107–108.

**Efthymios Costa**, Ioanna Papatsouma, Angelos Markos (2024). "A Deterministic Information Bottleneck Method for Clustering Mixed-Type Data". *arXiv preprint arXiv:2407.03389*.

**Efthymios Costa**, Ioanna Papatsouma, Angelos Markos (2022). "Benchmarking distance-based partitioning methods for mixed-type data". In: *Advances in Data Analysis and Classification* 17, pp. 701—724.

## FUNDING & AWARDS

---

### Poster Competition 1st Place & People's Choice Award

*Faculty of Natural Sciences Research Showcase 2025, Imperial College London*

### Poster Competition People's Choice Award

*Mathematics PhD Symposium 2025, Imperial College London*

### Best Postgraduate/Doctoral Paper Award

*6th Annual Conference of the Cyprus Statistical Society, 2025*

### Faculty of Natural Sciences Prize for Excellence in Teaching

*Highly commended nomination for excellence in teaching and learning, 2025, Imperial College London*

### MSc in Statistics Imperial College London Ambassador

*Honorary title in recognition of outstanding contributions to the running and delivery of the programme*

### Poster Competition Runner-up & People's Choice Award

*Mathematics PhD Symposium 2022, Imperial College London*

### PhD Studentship

*EPSRC Centre for Doctoral Training (CDT) in Statistics and Machine Learning (StatML)*

### Warner Prize

*Prize awarded for research project development in the MSc in Statistics at Imperial College London - Class of 2021*

### Outstanding Academic Representation Network Team Award

*Student volunteer award - Imperial College Union Student Choice Awards 2020*

## QUALIFICATIONS

---

### Associate Fellow of Advance HE (AFHEA)

*Awarded on February 2024, Imperial College London*

## CONFERENCES, WORKSHOPS & SEMINARS

---

### 19th International Conference on Computational and Methodological Statistics

*13 - 15 Dec 2025*

*Session Title: "Clustering of heterogeneous data"*

*Session Organiser/Chair*

### Joint meeting of the Italian and Dutch/Flemish Classification Societies (CLADAG-VOC 2025)

*8 - 10 Sep 2025*

*Presentation title: "Cluster Analysis from an Information-Theoretic Viewpoint"*

*Invited Speaker*

### International Conference on Robust Statistics (ICORS 2025)

*19 - 23 May 2025*

*Presentation title: "A novel framework for quantifying nominal outlyingness"*

*Invited Speaker*

### 18th International Conference on Computational and Methodological Statistics

*14 - 16 Dec 2024*

*Presentation title: "A Deterministic Information Bottleneck Method for Clustering Mixed-Type Data"*

*Presenter*

### RSS Emerging Applications Section Meeting "Network and Clustering Analysis for Emerging Applications"

*13 Dec 2024*

*Presentation title: "A Deterministic Information Bottleneck Method for Clustering Mixed-Type Data"*

*Invited Speaker*

### UKRI EPSRC Mathematical Sciences Team visit at Imperial College London

*9 Oct 2024*

*Poster title: "A Deterministic Information Bottleneck Method for Clustering Mixed-Type Data"*

*Presenter*

### 18th Conference of the International Federation of Classification Societies

*16 Jul 2024*

*Presentation title: "A Deterministic Information Bottleneck Method for Clustering Mixed-Type Data"*

*Presenter*

### Imperial Mathematics PhD Symposium

*12 Jun 2024*

*Poster title: "A Novel Approach to Outlier Detection for Mixed-Type Data"*

*Presenter*

### Department of Statistical Sciences "Paolo Fortunati", University of Bologna - Weekly Statistics Seminars

*1 Feb 2024*

*Presentation title: "Outlier detection for mixed-type data: A novel approach"*

*Presenter*

### 16th International Conference on Computational and Methodological Statistics

*16 - 18 Dec 2023*

*Poster title: "A novel approach to outlier detection for mixed-type data"*

*Presenter*

### Amazon Development Center Germany GmbH

*27 - 28 Mar 2023*

*Advanced training on Statistical Machine Learning and Computing for big-data analysis*

*Attendee*

*Poster title: "Benchmarking distance-based partitioning methods for mixed-type data"*

*Presenter*

<b>Faculty of Natural Sciences Research Showcase 2022</b>	<i>29 Sep 2022</i>
<i>Presentation title: "Clustering mixed-type data: Which method to choose?"</i>	Presenter
<b>Bocconi Summer School in Advanced Statistics &amp; Probability</b>	<i>11 - 22 Jul 2022</i>
<i>Topic: Random Structures and Combinational Statistics</i>	Attendee
<b>Imperial Mathematics PhD Symposium</b>	<i>16 Jun 2022</i>
<i>Poster title: "A full factorial benchmarking study of non-parametric partitioning methods for mixed-type data"</i>	Presenter

---

## TEACHING EXPERIENCE

### **BSc/MSci Mathematics Final Year Project Supervisor**

Imperial College London, Department of Mathematics (Academic Year 2025-2026)

### **MSc Statistics Dissertation Supervisor**

Imperial College London, Department of Mathematics (Academic Year 2025-2026)

### **Senior Graduate Teaching Assistant**

*Probability and Statistics* - BSc/MSci Mathematics Year 1 (Autumn 2024)

*Probability for Statistics* - BSc/MSci Mathematics Year 2 (Autumn 2022)

Imperial College London, Department of Mathematics

*Statistics* - MEng Mechanical Engineering Year 3/4 (Autumn 2023)

Imperial College London, Department of Mechanical Engineering

### **Graduate Teaching Assistant**

*Data Science* - MSc Statistics (Spring 2022)

*Computational Statistics* - MSc Statistics (Autumn 2022, 2023 & 2024)

*Unsupervised Learning* - MSc Machine Learning and Data Science (Autumn 2022)

*Deep Learning, Supervised Learning* - MSc Machine Learning and Data Science (Spring 2023)

*Machine Learning* - MSc Statistics (Spring 2023)

*Introduction to Statistical Learning, Statistical Modelling II* - BSc/MSci Mathematics Year 3/4 (Spring 2023)

*Statistical Modelling I* - BSc/MSci Mathematics Year 2 (Spring 2023)

Imperial College London, Department of Mathematics

---

## OUTREACH

Mentor for Imperial College London Mathematics School Research Project (2024 - 2025)

---

## EDITORIAL & PEER-REVIEW SERVICE

### **Reviewer**

*Advances in Data Analysis and Classification*

*International Statistical Review*

*Journal of Classification*

---

## MEMBERSHIPS

Fellow of the Royal Statistics Society (RSS)

Member of the British Data Science Society (BDSS)

Member of the Cyprus Statistics Society (CSS)

---

## SKILLS

**Programming/IT:** R (Advanced/Preferred), Python (Intermediate), C++ (Basic), MATLAB (Basic), L<sup>A</sup>T<sub>E</sub>X, HPC projects on compute clusters, Office 365

**Languages:** Greek (native), English (fluent), French (proficient)