

# Roberto Sánchez Navarro

Data Scientist - Data Analyst



 robertosanchezn

 robertosanchezn

## About me

Data Scientist looking for real-world challenges. Aspiring **Data Analyst** and **R/Python Developer** at CON&SEL15.

I'm curious about the data, so I often take the extra step until it makes sense. Then, I like presenting it in a way that tells a story.

I'm continuously under training, learning new technologies every week, and when I pick one up, I enjoy passing that knowledge along to my peers.

## Areas of specialization

Data Science • Data Analysis  
• Bioinformatics • Big Data

## Interests



Running • Cycling • Hiking

## References

**in** • Jonathan Rodriguez  
• Senior Software Engineer  
• Minsait


**in** • Morten Kam  
Dahl Dueholm • Associate  
Professor • Aalborg University

## EXPERIENCE

- 2022–2023 **PhD Candidate**  
AALBORG UNIVERSITY • Denmark   
Development of a python application for the detection and analysis of exopolysaccharide gene clusters in bacterial genomes and metagenomes.
- 2021–2022 **Research Assistant**  
AALBORG UNIVERSITY • Denmark   
Genome mining and analysis of metagenome-assembled genomes from wastewater treatment plants.



## EDUCATION

- 2023–2024 **Big Data Analysis**  
MSc. • University of Murcia 
- 2019–2021 **Biotechnology**  
MSc. • Aalborg University 
- 2015–2019 **Biotechnology**  
BSc. • University of Murcia 

## PROGRAMMING

<b>R</b>	4 years	<div></div>
<b>python</b>	3 years	<div></div>
<b>bash</b>	3 years	<div></div>
<b>LaTeX</b>	4 years	<div></div>
<b>SQL</b>	1 year	<div></div>

## TECHNICAL SKILLS

<b>Databases</b>	MySQL • MongoDB • HBase • Neo4j
<b>Data Wrangling</b>	pandas • dplyr • tidyr • purrr
<b>Visualization</b>	matplotlib • plotly • ggplot2 • ggiraph • Shiny
<b>Reporting</b>	R Markdown • jupyter • Google colab • Quarto
<b>Big Data</b>	Hadoop • slurm • Spark • Snakemake
<b>Office Tools</b>	Overleaf • Excel • PowerPoint • Slidy
<b>DB Clients</b>	sqlalchemy • pymongo • dbplyr • happybase
<b>Machine Learning</b>	caret • e1071 • Keras • scikit-learn • PyTorch (basic)
<b>Development</b>	conda/mamba • Docker • Git
<b>Scraping</b>	Scrapy • BeautifulSoup • rvest
<b>Other</b>	regex • API integration (Requests & http)

## LANGUAGES



<b>English</b>	C1 (Advanced) + full professional proficiency
<b>Spanish</b>	Mother tongue
<b>French</b>	B1 Certificate
<b>Danish</b>	Beginner

## PUBLICATIONS

**Sánchez-Navarro** et. al, 2022  
Long-Read Metagenome-Assembled Genomes Improve Identification of Novel Complete Biosynthetic Gene Clusters in a Complex Microbial Activated Sludge Ecosystem. *mSystems*. 10.1128/msystems.00632-22

## CODING COURSES

2019	DataCamp Career Track: <b>Data Scientist with R</b>
2021	DataCamp Career Track: <b>Python Programmer</b>

Roberto Sánchez Navarro  robertosan97@gmail.com  Spain  +34 644 902 167