

ANGELA CASTILLO-GILL

Data Scientist

[LinkedIn](#) • Phone: +44 7494 098822 • E-mail: acastillogill@gmail.com • [Twitter](#) • [Data Science Portfolio](#)

We all want to save time, space, reduce waste, and we want to earn more, increase brand sales and profit. That's why I'm passionate about **machine learning** and **data science** - it allows us to overcome human biases and make better decisions to reach our targets. My degree in industrial engineering has given me the skills to use new techniques in the field of marketing, consumer research, predictive analytics, and operations research. Since new tools and techniques are constantly appearing, to stay up to date, every month I receive a different data-related online certification.

DATA SCIENCE SKILLSET

Languages	Machine learning	Data Viz tools	Libraries	Statistics	Character traits
-R	-Linear regression	-Excel	-Dplyr	-Linear algebra	-Numerate
-SQL	-Logistic regression	-Google Charts	-mlr	-Multivariate calculus	-Eye for detail
-Python	-PCA	-ggplot2	-Caret	-Probability distributions	-Team player
-Microsoft Excel	-Clustering	-Tableau	-Gbm	-Hypothesis testing	-Quick learner
	-CART	-Shiny	-XGBoost	-Time series analysis	-Strong communicator
	-Random Forests		-pandas	-Forecasting	
	-k-NN		-seaborn	-A/B testing	

PROFESSIONAL EXPERIENCE

MANGO SOLUTIONS London, United Kingdom Data Scientist

Feb 2019 - Present

- Built a Shiny dashboard for one of the largest insurance and resinsurance markets in the world, this entailed the following day-to-day activities:
 - Wrangling large databases, more than 1.5m rows.
 - Creating SQL queries to optimise dashboard running times.
 - Writing efficient R code to minimise computational resources.
- Currently developing a tool to improve the calculation of insurance premiums for properties around the world integrating machine learning algorithms, proprietary knowledge, and open-source data.

UNIVERSITAT AUTÒNOMA DE BARCELONA Barcelona, Spain Data Scientist | Researcher

Feb 2017 - Jan 2019

- Models built and submitted to peer-reviewed publication on 27/11/18 to *Ecological Economics*: two multiple linear regression (MLR) and one time series forecast.
 - MLR: One to predict property sales prices and the other to predict number of trees according to spatial location.
 - Time series forecast: To predict the number of invasive species detected in the Iberian Peninsula.
- Performed end-to-end data analysis which involved database design, data cleaning, visualisation, and report writing to communicate findings to non-technical colleagues.
- Built a spatial-temporal model to simulate an international law and inform policy makers.
- Proposed and fostered collaboration with the data science team of one of the biggest online real-estate companies in Spain, Idealista SA in order to use their data for research.
- Collaborated with non-technical colleagues including government stakeholders to create an ethical and well-informed model.

- Communicated data insights at several international public-speaking engagements to promote the research institute.

INFORMA AGRIBUSINESS INTELLIGENCE

London, United Kingdom

Data Analysis Consultant

Apr 2016 – Mar 2017

- Built time series visualisations to show the trend in commodities of interest.
- Performed end-to-end data analysis which involved database design, data cleaning, visualisation, and report writing on the subjects:

1. Quinoa Markets 2016 – Trends and Opportunities
2. Colombian Agribusiness after FARC peace deals
3. Delivering the Sustainable Development Goals (SDGs) in Agrifood Companies 2017

UNIVERSIDAD SERGIO ARBOLEDA

Bogotá, Colombia

Analysis of Cleaner Production Lecturer

Aug 2016 - Dec 2016

- Designed the module Best Practices in Cleaner Production geared towards last year Industrial Engineering Students.
- Motivated the implementation of Best Practices in Cleaner Production in small and medium sized enterprises (SME's) in Bogotá, Colombia.

HRA FOOD AND DRINK

Torquay, United Kingdom

Data Analysis Intern

Oct 2015 - Jan 2016

- Built a database of all products in the relevant industry from a different number of sources.
- Analysed every aspect of the British flavoured milk industry including main producers, brands, formulas, consumer trends, packaging and distribution.
- Suggested innovation strategies for the dairy industry and niche food markets in the UK.
- Wrote reports for the British flavoured-milk industry that included consolidated trends and suggested growth areas.

EDUCATION

UNIVERSITY OF EXETER, Master of Science in Food Security & Sustainable Agriculture
Exeter, United Kingdom – Grade: Distinction Award

2014 - 2015

Research project: [*Published in journal Philosophical Transaction of the Royal Society.*](#)

UNIVERSIDAD DE LOS ANDES, Bachelor of Science in Industrial Engineering
Bogotá, Colombia – Grade: 4.02/5 (top 10% of cohort)

2010 - 2014

Key modules: Probability and Statistics, Optimization, Finances, Logistics, Organizations.

GRANTS AND AWARDS

María de Maetzu, Spain

2017

Spanish Ministry of Economy and Competitiveness (MINECO) – *Ten scholarships granted each year.*
 Awarded a scholarship worth €50,000 to pursue a María de Maetzu PhD. 'María de Maeztu' Unit of Excellence is the highest institutional recognition of scientific research in Spain.

IBM, United Kingdom

2015

IBM Best Student Recognition Event 2015 – *Two students selected from each country per year.*
 Awarded a place at the IBM Best Student Recognition Event in Amsterdam, Netherlands.

Ayudar Scholarship – *Only two scholarships granted per year.*

Awarded £30,000 to fund tuition and maintenance fees during postgraduate studies.

LANGUAGE SKILLS

English (native), Spanish (native) and French (intermediate).

APPENDIX – ONLINE CERTIFICATIONS

R

Title	Issued by	Completion date
Introduction to R	DataCamp	January 2018
Introduction to the Tidyverse	DataCamp	January 2018
Working with Geospatial Data in R	DataCamp	February 2018
Intermediate R	DataCamp	February 2018
Cleaning Data in R	DataCamp	February 2018
Data Manipulation in R with dplyr	DataCamp	March 2018
Data Visualization with ggplot2 (Part 1)	DataCamp	March 2018
Data Visualization with ggplot2 (Part 2)	DataCamp	March 2018
Data Visualization with ggplot2 (Part 3)	DataCamp	April 2018
Spatial Analysis in R with sf and raster	DataCamp	April 2018
Reporting with R Markdown	DataCamp	April 2018
Exploratory Data Analysis	DataCamp	April 2018
Joining Data in R with dplyr	DataCamp	May 2018
Introduction to Data	DataCamp	May 2018
Writing Functions in R	DataCamp	May 2018
Working with Dates and Times in R	DataCamp	May 2018
Correlation and Regression	DataCamp	June 2018
Exploratory Data Analysis in R: Case Study	DataCamp	June 2018
Importing Data in R (Part 1)	DataCamp	July 2018
Importing Data in R (Part 2)	DataCamp	July 2018
Importing & Cleaning Data in R: Case Studies	DataCamp	August 2018
Sentiment Analysis in R: The Tidy Way	DataCamp	August 2018
Cluster Analysis in R	DataCamp	September 2018
Unsupervised Learning in R	DataCamp	September 2018
Supervised Learning in R: Classification	DataCamp	September 2018
Machine Learning with Tree-Based Models in R	DataCamp	September 2018
Manipulating Time Series Data in R with xts & zoo	DataCamp	November 2018
Introduction to Time Series Analysis	DataCamp	November 2018
ARIMA Modeling with R	DataCamp	December 2018
Forecasting Using R	DataCamp	December 2018
Machine Learning in the Tidyverse	DataCamp	December 2018
Developing R Packages	DataCamp	January 2019
A/B Testing in R	DataCamp	February 2019
Building Web Applications in R with Shiny	DataCamp	March 2019

SQL

Intro to SQL for Data Science	DataCamp	June 2018
Joining Data in SQL	DataCamp	June 2018

Python

Introduction to Python	DataCamp	January 2019
Python for R Users	DataCamp	January 2019

Intermediate Python for Data Science	DataCamp	February 2019
Python Data Science Toolbox (Part 1)	DataCamp	February 2019
Python Data Science Toolbox (Part 2)	DataCamp	March 2019
Importing Data in Python (Part 1)	DataCamp	April 2019
Importing Data in Python (Part 2)	DataCamp	April 2019
Cleaning Data in Python	DataCamp	May 2019
Pandas Foundations	DataCamp	June 2019

Business Analytics

Operations Analytics	Coursera	November 2018
Customer Analytics	Coursera	December 2018
People Analytics	Coursera	January 2019

Google Cloud Platform

Google Cloud Platform Big Data and Machine Learning Fundamentals	Coursera	May 2019
--	----------	----------