

# Juan Jose Carin

Data Scientist

Mountain View, CA 94041  
650-336-4590 | [juanjose.carin@gmail.com](mailto:juanjose.carin@gmail.com)  
[linkedin.com/in/juanjosecarin](https://www.linkedin.com/in/juanjosecarin) | [juanjocarin.github.io](https://juanjocarin.github.io)

## Professional Profile

Passionate about data analysis and experiments, mainly focused on user behavior, experience, and engagement, with a solid background in data science and statistics, and extensive experience using data insights to drive business growth.

## Education

2016	<b>University of California, Berkeley</b>	<b>Master of Information and Data Science</b>	GPA: 3.93
	<i>Relevant courses:</i>		
	<ul style="list-style-type: none"><li>Machine Learning</li><li>Machine Learning at Scale</li><li>Storing and Retrieving Data</li></ul>	<ul style="list-style-type: none"><li>Field Experiments</li><li>Applied Regression and Time Series Analysis</li><li>Exploring and Analyzing Data</li></ul>	<ul style="list-style-type: none"><li>Data Visualization and Communication</li><li>Research Design and Applications for Data Analysis</li></ul>
2014	<b>Universidad Politécnica de Madrid</b>	<b>M.S. in Statistical and Computational Information Processing</b>	GPA: 3.69
	<i>Relevant courses:</i>		
	<ul style="list-style-type: none"><li>Data Mining</li><li>Multivariate Analysis</li><li>Time Series</li></ul>	<ul style="list-style-type: none"><li>Neural Networks and Statistical Learning</li><li>Regression and Prediction Methods</li><li>Optimization Techniques</li></ul>	<ul style="list-style-type: none"><li>Monte Carlo Techniques</li><li>Numerical Methods in Finance</li><li>Stochastic Models in Finance</li><li>Bayesian Networks</li></ul>
2005	<b>Universidad Politécnica de Madrid</b>	<b>M.S. in Telecommunication Engineering</b>	GPA: 3.03
	<i>Focus Area:</i>	Radio communication systems (radar and mobile).	
	<i>Fellowship:</i>	First year at University, due to Honors obtained last year at high school.	

## Skills

	<u>Programming / Statistics</u>	<u>Big Data</u>	<u>Visualization</u>	<u>Others</u>
Proficient:	<i>R, Python, SQL</i>	<i>Hadoop, Hive, MrJob</i>	<i>Tableau</i>	<i>Git, AWS</i>
Intermediate:	<i>SPSS, SAS, Matlab</i>	<i>Spark, Storm</i>		<i>Bash</i>
Basic:	<i>EViews, Demetra+</i>		<i>D3.js</i>	<i>Gephi, Neo4j, QGIS</i>

## Experience

### DATA SCIENCE

Jan. 2016 – Mar. 2016	<b>Data Scientist</b> CONENTO	Madrid, Spain (working remotely)
	<ul style="list-style-type: none"><li>Designed and implemented the ETL pipeline for a predictive model of traffic on the main roads in eastern Spain (a project for the Spanish government).</li><li>Automated scripts in <i>R</i> to extract, transform, clean (incl. anomaly detection), and load into <i>MySQL</i> data from multiple data sources: road traffic sensors, accidents, road works, weather.</li></ul>	
Jun. 2014 – Sep. 2014	<b>Data Scientist</b> CONENTO	Madrid, Spain
	<ul style="list-style-type: none"><li>Designed an experiment for Google Spain (conducted in October 2014) to measure the impact of YouTube ads on the sales of a car manufacturer's dealer network.</li><li>A matched-pair, cluster-randomized design, which involved selecting the test and control groups from a sample of 50+ cities in Spain (where geo-targeted ads were possible) based on their sales-wise similarity over time, using wavelets (and <i>R</i>).</li></ul>	

### MANAGEMENT – SALES (Electrical Eng.)

Feb. 2009 – Aug. 2013	<b>Head of Sales, Spain &amp; Portugal – Test &amp; Measurement dept.</b> YOKOGAWA	Madrid, Spain
	<ul style="list-style-type: none"><li>Applied analysis of sales and market trends to decide the direction of the department.</li><li>Led a team of 7 people.</li></ul>	

# Juan Jose Carin

## Data Scientist

Mountain View, CA 94041  
650-336-4590 | [juanjose.carin@gmail.com](mailto:juanjose.carin@gmail.com)  
[linkedin.com/in/juanjosecarin](https://www.linkedin.com/in/juanjosecarin) | [juanjocarin.github.io](https://juanjocarin.github.io)

- Increased revenue by 6.3%, gross profit by 4.2%, and operating income by 146%, and achieved a 30% ratio of new customers (3x growth), by entering new markets and improving customer service and training.

### SALES (Electrical Eng. & Telecom.)

- Apr. 2008 – Jan. 2009 **Sales Engineer – Test & Measurement dept.**  
YOKOGAWA Madrid, Spain
- Promoted to head of sales after 5 months leading the sales team.
- Sep. 2004 – Mar. 2008 **Sales & Application Engineer**  
AYSCOM Madrid, Spain
- Exceeded sales target every year from 2005 to 2007 (achieved 60% of the target in the first 3 months of 2008).

### EDUCATION

- Jul. 2002 – Jun. 2004 **Tutor of Differential & Integral Calculus, Physics, and Digital Electronic Circuits**  
ACADEMIA UNIVERSITARIA Madrid, Spain
- Highest-rated professor in student surveys, in 4 of the 6 terms.
  - Increased ratio of students passing the course by 25%.

### Projects

See [juanjocarin.github.io](https://juanjocarin.github.io) for additional information

- 2016 **SmartCam**  
Capstone *Python, OpenCV, TensorFlow, AWS (EC2, S3, DynamoDB)*  
A scalable cloud-based video monitoring system that features motion detection, face counting, and image recognition.
- 2015 **Implementation of the Shortest Path and PageRank algorithms with the Wikipedia graph dataset**  
*Machine Learning at Scale Hadoop MrJob, Python, AWS EC2, AWS S3*  
Using a graph dataset of almost half a million nodes.
- 2015 **Forest cover type prediction**  
*Machine Learning Python, Scikit-Learn, Matplotlib*  
A Kaggle competition: predictions of the predominant kind of tree cover, from strictly cartographic variables such as elevation and soil type, using random forests, SVMs, kNNs, Naive Bayes, Gradient Descent, GMMs, ...
- 2015 **Redefining the job search process**  
*Storing and Retrieving Data Hadoop HDFS, Hive, Spark, Python, AWS EC2, Tableau*  
A pipeline that combines data from Indeed API and the U.S. Census Bureau to select the best locations for data scientists based on the number of job postings, housing cost, etc.
- 2015 **A fresh perspective on Citi Bike**  
*Data Visualization and Communication Tableau, SQLite*  
An interactive website to visualize NYC Citi Bike bicycle sharing service.
- 2015 **Investigating the effect of competition on the ability to solve arithmetic problems**  
*Field Experiments R*  
A randomized controlled trial in which 300+ participants were assigned to a control group or one of two test groups to evaluate the effect of competition (being compared to no one or someone better or worse).
- 2014 **Prediction of customer churn for a mobile network carrier**  
*Data Mining SAS*  
Predictions from a sample of 45,000+ customers, using tree decisions, logistic regression, and neural networks.
- 2014 **Different models of Harmonized Index of Consumer Prices (HICP) in Spain**  
*Time Series SPSS, Demetra+*  
Forecasts based on exponential smoothing, ARIMA, and transfer function (using petrol price as independent variable) models.