MANUEL RUEDA

New York, NY, USA 10033

manuel.rueda@columbia.edu

(929) 235-3220

EDUCATION

M.S. in Data Science at Columbia University (Dec 2016, GPA 3.65)

- **Courses:** Algorithms, Machine Learning (ML), Bayesian Models for ML, Natural Language Processing (NLP), Computational Models of Social Meaning, Data Visualization, Distributed Computing, GIS & Spatial Analysis.
- **Capstone Project:** Sponsored by Goldman Sachs. Used natural language processing and deep learning tools to programmatically quantify changes on 10-K company reports.

B.Sc. in Economics at Tec de Monterrey - ITESM (Dec 2009, GPA 3.50)

- **Courses:** Econometrics, Time Series, Multivariate Analysis, Statistics, Mathematics for Engineering, Linear Algebra.

Certified Financial Risk Manager (Jan 2014, GARP)

- **Curriculum:** Quantitative Analysis, Financial Markets and Products, Valuation and Risk Models, Risk Management, Market Risk, Credit Risk, Operational Risk, Current Issues on Financial Markets.

PROFESSIONAL EXPERIENCE

Data Science Intern at Kora Management LP. (June 2016 – Dec 2016, NY)

Kora is a global hedge fund that leverages on data science to make long term investment decisions.

- Programmatically scrapped unstructured transactional data from e-commerce websites and analyzed it with *Spark* (*Scrapy*, *AWS EMR*).
- Applied feature engineering on this data to build predictive models for companies' performance (*python, Spark*).
- Performed statistical analysis and inference on survey data (*python*).

Data Science Intern at Trinnacle Capital Management (Feb 2016 – May 2016, NY)

Trinnacle is a quantitative hedge fund that uses unconventional data sources to discover investment opportunities.

- Created and administered databases with financial and market big data (*Hadoop, MySQL, AWS*).
- Built a model to identify short-term investment opportunities based on earnings announcements. Using historical minute-frequency data for the last 10 years, I applied backtesting (i.e. time series cross-validation) to find the best model parameters (*R*).
- Created a model that uses mobile GPS data to estimate the number of customers visiting a target business, and then predict its revenue. The first test case predicted with accuracy superior to Bloomberg's estimates *(R, Shiny)*.

Manager of the Market Data Team at MSCI (Jan 2010 – Aug 2015, MX)

Asides from maintaining the MSCI global indices, the company provides models for risk analysis and portfolio optimization.

- Built and managed the Market Data Productions team**,** leading 3 analysts.
- Worked with Research, Engineering and Project Management to build risk models and incorporate clients' suggestions.
- Developed interactive tools and dashboards for analysts to analyze thousands of time series efficiently *(SQL, R, Python, SpotFire)*.
- Developed scripts to cross-validate and stress test risk models, generating automated reports for them (*SQL, R*).

TA for Columbia's edX Courses (Dec 2015 – Feb 2016, NY)

- Statistical Thinking for Data Science and Analytics.
- Machine Learning for Data Science and Analytics.

DATA PROJECTS (http://masta-g3.github.io)

Linear Content Blog (Ongoing)**:** Personal blog where I post experiments on Machine Learning and NLP (*python*).

Sentiment Analysis for Stock Prediction (Fall 2016)**:** Analyzing million of Amazon's reviews on brands and products (*python, Spark*).

Summer Networks (Summer 2016)**:** Several implementations of a neural network that can write short poems (*numpy, theano*).

MoMA Through Time (Feb 2015)**:** HTML visualization of exhibitions patterns on the museum; awarded best of the hackathon (*R, D3*).

GENERAL SKILLS

Programming: R & Python.

Visualization: Shiny, SpotFire, D3 & P5.

DB & Cloud Computing: SQL, AWS & Spark.

GIS Analysis: QGIS, GeoDa & CartoDB.

Web: HTML, CSS & JS.