MANUEL RUEDA   New York, NY, USA 10033

manuel.rueda@columbia.edu   masta-g3.github.io

EDUCATION

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

- Algorithms, Machine Learning, Bayesian ML, NLP, Data Visualization, GIS & Spatial Analysis, Distributed Computing.
- Capstone Project (Goldman-Sachs): Use text similarity measures and deep learning tools to quantify changes on financial reports from the SEC.

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

- Econometrics & Time Series, Multivariate Analysis, Statistics, Mathematics for Engineering, Linear Algebra.

Certified Risk Manager (Jan 2014)

- Certified Financial Risk Manager (FRM) by the Global Association of Risk Professionals (GARP).

PROFESSIONAL EXPERIENCE

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

Kora is a global hedge fund that leverages on data science to make long term investment decisions.
- Scrapped unstructured data from the web, aggregated and analyzed it with *Spark* (*Scrapy*, *AWS EMR*).
- Applied feature engineering to build unsupervised learning models and track companies' performance (*python, Spark-TS*).
- Performed statistical analysis and inference on survey data (*python*).
- Implemented a text polarity tweet classifier that collects the data, analyzes it and displays the results as time series on an interactive dashboard (*python, MongoDB, R, Shiny*).

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

Trinnacle is a quantitative hedge fund that uses unconventional data sources to find market opportunities.
- Created and administered HDFS and SQL databases with financial and market big data (*Qubole, MySQL, AWS*).
- Built a model to identify short-term investment opportunities using historical minute-frequency data for the last 10 years (*R*).
- Created a model and dashboard that uses mobile GPS data to predict quarterly revenues for companies *(R, Shiny)*.

Senior Associate at MSCI (Jan 2010 – Aug 2015, MX)

Asides from maintaining the MSCI indices, the company provides analytical models for risk management (RiskMetrics).
- Built and managed the Market Data Productions team**,** leading 3 analysts.
- Coordinated projects with the global Research and Project Management teams.
- Developed interactive dashboards used by analysts to perform daily QA on >10,000 time series *(R, SpotFire)*.
- Developed automated reporting scripts in R for risk models validation (*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 SCIENCE PROJECTS (masta-g3.github.io)

Geospatial Analysis of Gentrification in NYC (Dec 2015): Using social and economical geo-data (*R, QGIS, CartoDB*).

CULPA Sentiment (Feb 2015): Sentiment analysis and visualization of Columbia's academic course ratings (*Python, HTML, D3*).

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

Summer Networks (Summer 2016): Several tutorials and experiments on Neural Networks, on my attempt to build a RNN that can write short poems (*python, theano*).

GENERAL SKILLS

Visualization: SpotFire, Shiny, D3 & P5.js.   **GIS Analysis:** QGIS, GeoDa & CartoDB.

Programming: R & Python.   **Web:** HTML, CSS & JS.

DB & Cloud Computing: SQL, AWS & Spark.