

CRUTAIN Yves

17 Henschman street
W12 0BW London, UK
Tel: +44 (0)7546605818
yves.crutain@gmail.com

Data scientist
R, SQL, Data mining
Six years' experience
data-laborer.eu

PROFILE

- ✓ Five years of working experience in Marketing Analytics and R&D.
- ✓ Career driven, self-learner, well organized and energetic.
- ✓ Strong technical background, knowledge of all major statistical methods.
- ✓ Passionate about Data Mining and Computer Science.
- ✓ Engaged in data science communities around London.
- ✓ Follower of good working policies.
- ✓ Genuine curiosity for other people, cultures, and countries. Blogger.

PROFESSIONAL EXPERIENCES

MODEL CITIZENS (2 years)

- Data scientist in a marketing environment.

CETELEM / BNP PARIBAS (1 year)

- Basel II developer on statistical projects.

KEYRUS: Consultant for key accounts

- GDF-SUEZ - CEEME (< 1 year)
 - ❑ Developer/ analyst in an econometric department.
- EDF (<3 years)
 - ❑ Insight analyst in a marketing environment.
 - ❑ Analyst in a research environment.

EXPERTISE FIELD

- **Statistics and probabilities, machine learning, time series**
- **R**. Packages: openCPU, shiny, ggplot2, tidyverse, data.table, h2o
- SAS v9.3, Base, Stat, Macro, SQL, Connect to DB2, Connect to Oracle, IML, SEG

EDUCATION

- **2012**: Coursera classes on machine learning and econometrics
- **2010**: Master's degree: mathematics for computer graphics and statistics.
- **2008**: License mathematics (University of Burgundy – Dijon)

OTHER SKILLS

- French, mother tongue
- English: professional abilities & Toastmaster member
- Computing language **SQL, R, python**(scikit-learn, numpy), C++, latex, HTML, CSS
- Software Matlab/Scilab, Courbotree, Khiops, Kiphren, H2O
- Google technology (Google Refine, Google Visualization API, etc.)
- European driving license (B)

PROFESSIONAL EXPERIENCES, DETAILS:

October 2014 - Current position: Data Scientist

MODEL CITIZENS

- Creation of SQL views and dashboards for key accounts.
- Creation and development of machine learning based tools: Recommendation engine, REST API (openCPU).
- Ad hoc studies.

Keyword: R (RODBC, data.table, shiny, ggplot2, glmnet, etc.), SQL server, HTML, CSS, OpenCPU, Recommendation Systems, Regression

September 2013 - October 2014: Basel II Development

CETELEM / BNP PARIBAS

- SAS programming in big data environment, including data management and statistical analyses.
- Optimisation of oracle queries.
- Steering of three SAS consultant.

Keyword: SAS (SEG, base, stat, graph, macro, SQL, connect to oracle, SAS Stored Processes), hash code, Microsoft office add-on, Excel, Bale II(LGD, PD, EAD), oracle, statistical tests.

February 2013: Trainer, R Software

October 2012 - June 2013: Modeling of the energy market from 2012 to 2030

GDF - SUEZ Center of Expertise in Economic Modeling and Studies (CEEME)

- Programming in SAS the forecast to 2030 of the world energy consumption.
- Implementation of a datamart.
- Ad hoc studies (climate correction, time series smoothing, etc.)

Keyword: SAS(IML, macro), data management, AIE, Enerdata and GDF-SUEZ data.

February 2012 – September 2012: Study of the consumption of a panel, under erasing condition EDF R&D Business innovation and market analysis of energy

- Management of a panel (data quality, data processing, survey selection).
- Comparison of regression method: lasso, PLS, smoothing, linear, ARIMA.
- **Redaction of a complete study on the panel's consumption, including sociological, temporal, weather and material influences.**

Keyword: SAS(Stat, connect to DB2), R, INSEE data, regression, temporal series, data intensive environment

July 2011 – December 2011: Study of the influence of a migration on the customer's bill

EDF BRANCH TRADE - Analysis and customer knowledge department

- Programming of two models of estimation of electric consumption.
- Analysis of the influence of the new estimation on the cash working capital. Prediction of the gain / loss by type of customer.

Keyword: SAS(Base, macro, connect to oracle), data management, SIMM, QE, descriptive statistics

March 2011 – Jun 2011: Study of multi-site customers in EDF database

EDF BRANCH TRADE - Analysis and customer knowledge department

- Population profiling.
- **Definition of specific vocabulary and appropriate metrics for the study.**
- **Classification / clustering and statistical inferences.**
- Production of a score still in use to contact people and propose them to coerce their multiple contracts into a single one.

Keyword: **SAS(Base, Stat), SIG, data intensive environment**

November 2010 – February 2011: Study of the influence of e-mail campaign on gas customer's comportment.

EDF BRANCH TRADE - Analysis and customer knowledge department

- Population profiling.
- Documentation redaction.

Keyword: **SAS, excel, descriptive statistics, QE**

October 2010: Clustering and classification of the enterprise on the french territory.

EDF R&D Business innovation and market analysis of energy

- Clustering of geographical data for GIS uses, using naives bayes predictor.

Keyword: **Bayesian probability, GIS, clustering/classification, Khiops, Khiphren**

May 2010 - September 2010: Study of reconstruction curve (internship)

EDF R&D Business innovation and market analysis of energy

- Development of experimental algorithm for time series coercion.

Keyword: **SAS(Base, macro, stat), R, data management, latex, courboscope**

December 2009 – April 2010: Clustering on INRIA's anonymous data (dissertation)

University of Burgundy

- Documentation on the SVM.
- Synthesis and redaction of the theoretical work.
- Use of an open source matlab package for practical work.

Keyword: **Matlab, SVM, latex**

SIDE PROJECTS

LONDON DATA SCIENCE COMMUNITY Hackathon:

December 2014: Health data hack, first prize.

January 2014: IoT hack, second prize.

March 2014: Hack the brain, second prize.

Member of the data scientist speaker club (toastmaster club).

My blog on data science: data-laborer.eu

My Github profile: github.com/YvesCR

My stackoverflow profile: stackoverflow.com/users/4911229/ycr