



# WHO IS THE DATA SCIENTIST THE MARKET IS SEARCHING FOR?

## CRYSTAL SYSTEM GROUP LTD.

**IN ORDER TO UNDERSTAND THE PROFILE OF THE DATA SCIENTIST THE MARKET** (by market we understand HR managers and recruiting specialists as the main agents in charge to hire the right professionals for business as well as the business leaders who define the requirements) is searching for; we analyzed 90% of the different job offers\* for the Data Scientist placed on such platforms as Glassdoor, LinkedIn and slb.dejobs.org (February – March 2020).

*\*These job offers are profiles of the Data Scientist searched by the companies that are leaders in different industries such as analytics, software and engineering, energy, consumer goods, banking, information technologies, tobacco.*

The goal was to understand the profile of the so called “Market Data Scientist” (the ideal candidate for the market) through the following 5 categories:

- 1) “Market Data Scientist” - Education requirements
- 2) “Market Data Scientist” - Tools and technologies capacity requirements
- 3) “Market Data Scientist” - Business acumen requirements
- 4) “Market Data Scientist” - Soft skills and foreign languages requirements
- 5) “Market Data Scientist” - Professional activities requirements

Having this information, we were able to define:

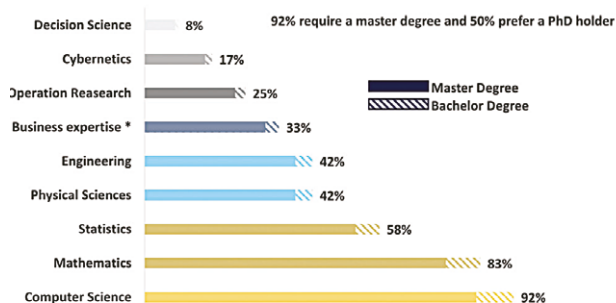
- WHO is the ideal Data Scientist the market is searching for;
- if the market “mission” to find the ideal candidate is possible;
- and how companies can apply Crystal Pragmatic Approach to BUILD and SELECT Data Scientist Talents.

### 1) “MARKET DATA SCIENTIST” – EDUCATION REQUIREMENTS

The bar chart (Figure 1) illustrates the education domains preferred by the HR for the Data Scientist candidate. It can be seen that the Computer Science Master degree is required by the 92% of the market searchers, and then followed by the Math and Statistics graduates (83% and 58% respectively).

It is worth noting that some of the employers in their announcements do not specify clearly the education domain of the Data Scientist searched for their business.

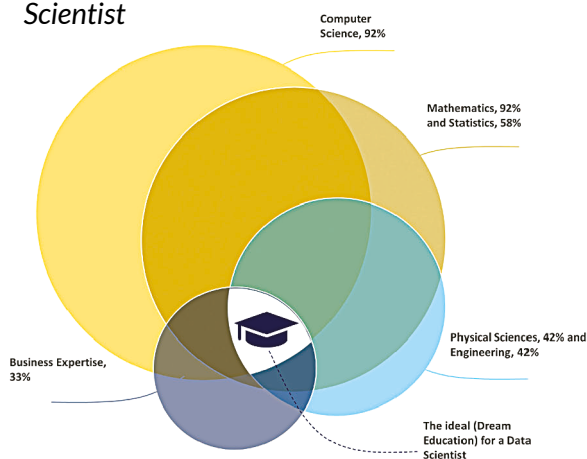
(Figure 1) Data Scientist Education Domain searched by the market



Such statements as “any quantitative field”, “other technical field”, “other quantitative/computational discipline”, “another field relevant to Data Science”, etc. may puzzle the potential work searchers and might be considered as a market confusion.

The diagram (Figure 2) shows how the education domains for the Data Science professional overlap and the size of the connecting part – the ideal education set for the Data Scientist.

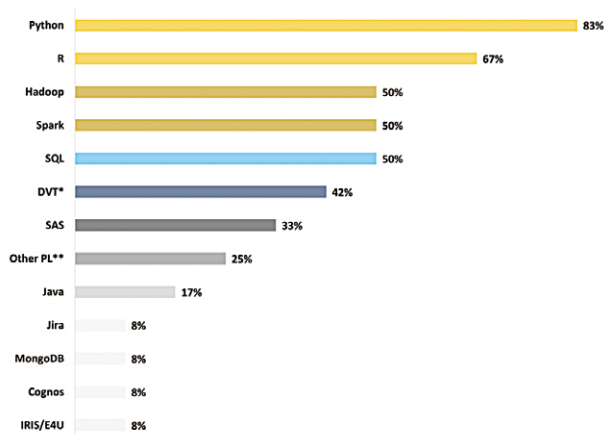
(Figure 2) The ideal education set of the Data Scientist



## 2) “MARKET DATA SCIENTIST” TOOLS AND TECHNOLOGIES CAPACITY

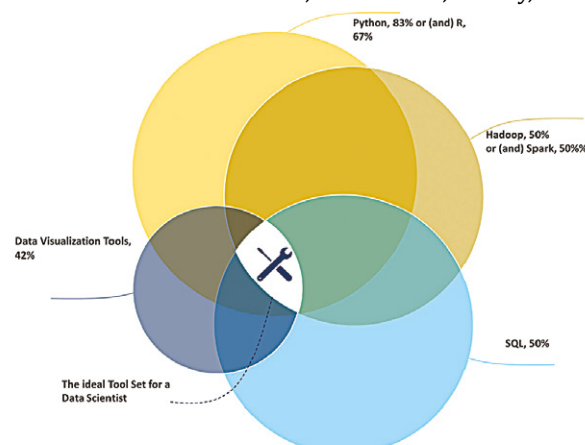
Summarizing the information regarding tools and technologies requirements (Figure 3), Python and R are obviously “champions” of search within 83% and 67% HR managers respectively, and it’s clear as these tools provide very neat means for data collection, modelling and visualization stages of a data Science Project.

Figure 3 Data Scientist tools and technology capacity market requirements



Remarkable, that in addition to the most popular tools and technologies represented in the charts, the market is searching also the candidate to have some knowledge in Machine Learning technologies such as Tensorflow, Theano, MXNet, PyTorch, and some of them – in in data mining software (e.g. Weka, Knime, Databricks, Tensorflow, Scikit-learn).

Some employers ask for experience with such Data Science libraries like Pandas, Scikit-Learn, NumPy, etc.



(Figure 4) The ideal tool set of the Data Scientist

The diagram (Figure 4) illustrates how the overlapping of the most popular tools and technologies required from the Data Science professional by the HRs leaders– the ideal tool/technology set for the Data Scientist.

## 3) “MARKET DATA SCIENTIST” BUSINESS ACUMEN REQUIREMENTS.

If to sum up the experience requirements for the Data Scientist (Figure 5), the ideal candidate should have 3-5 years’ experience in the industry or at least poses some business knowledge as he should provide C-suit with business insights and be able to communicate results to business community.

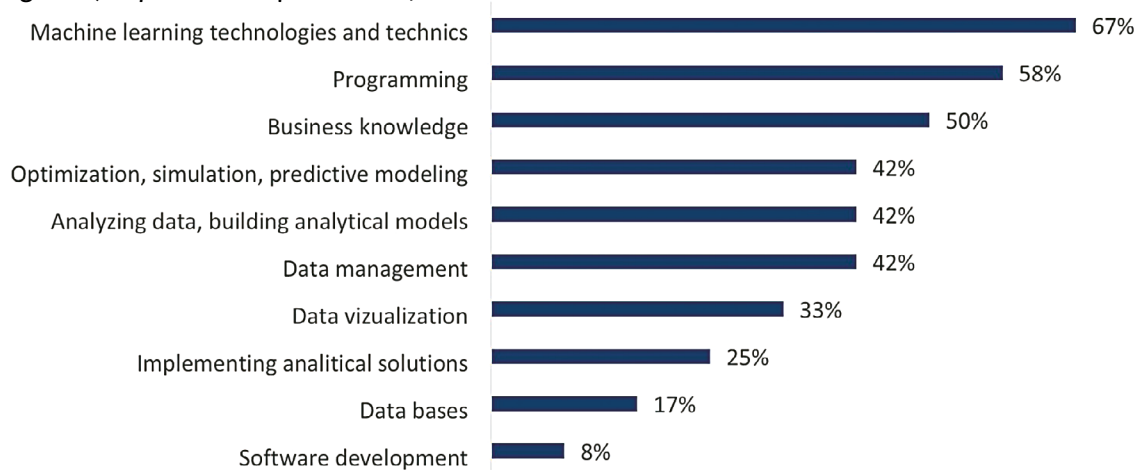
Business Acumen (as a possession of business knowledge in a specified area) is in the top 3 experience requirements for the Data Scientist searched by the market.

Meanwhile, the ideal candidate should have a working background by using ML technologies (required by 67% of the employers) and have sufficient proved skills in programming (needed by 58% of the companies).

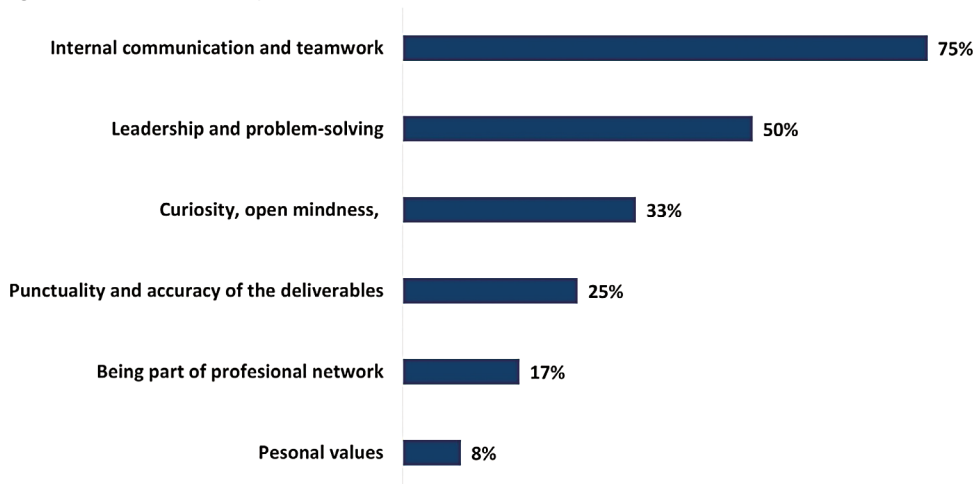
## 4) “MARKET DATA SCIENTIST” SOFT SKILLS AND FOREIGN LANGUAGES REQUIREMENTS

75% of the employers require the Data Scientist to be able to communicate effectively within internal departments as well as to be a team player, and 50% of them expect the candidate to have leader skills

(Figure 5) Experience requirements for the Data Scientist



(Figure 6) Soft skills requirements for the Data Scientist



and problem-solving mind-set (Figure 6). Moreover 100% of employers need a Data Science professional with excellent written and verbal English, and good understanding of a second language such as German, Italian, or French is preferred by 30% of the employers.

### 5) “MARKET DATA SCIENTIST” PROFESSIONAL ACTIVITIES REQUIREMENTS

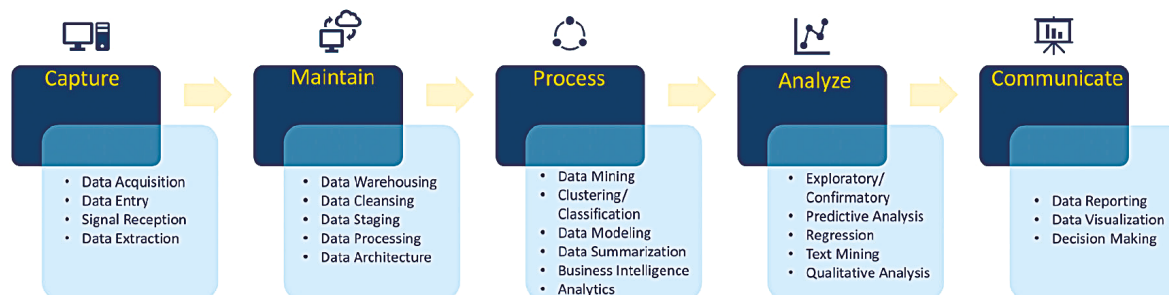
By professional activities requirements we understand the professional responsibilities and expectations from the Data Scientist employee. Analyzing the job offers, most of the employers require the Data Scientist to be able to master the full spectrum of the Data Science Life Cycle and possess

a level of flexibility and understanding to maximize returns at each phase of the process (Figure 7).

Summarizing the main requirements from the search categories from above, we can highlight the following key requirements from the Ideal Data Scientist candidate hunted by the HRs:

- Master degree in Computer Science, Math or Statistics (and preferably PhD)
- Strong knowledge and working experience (3-5 years) with Python and/or R, and knowledge of Machine Learning technologies and tools
- High level of business understanding (energy, construction, accountancy, banking, other businesses depending on the sector of employment), as the

(Figure 7) Soft skills requirements for the Data Scientist



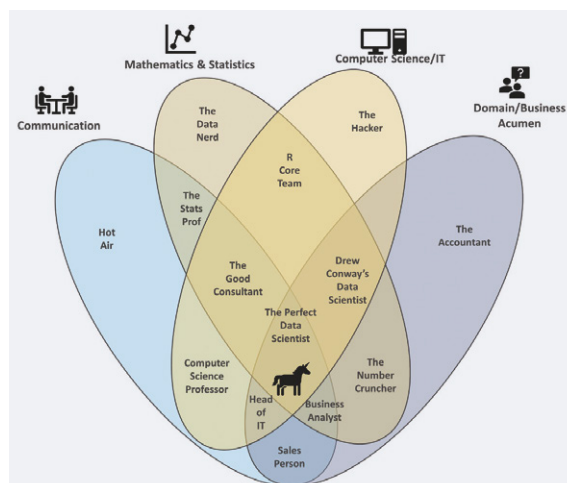
candidate should provide business managers with business insights and prepare data for strategic decision making

- English is a must speak language, and a second language such as French, Italian, Spanish or German is required for positions where candidate should communicate with foreign departments
- Teamwork, leadership and problem-solving mindset
- We can continue with the list, ... but the question that appears is do you really expect to find all these knowledge and skills in one person?

The overlapping part of the diagram (Figure 8) shows us the approximate connecting point of the category requirements from the Ideal Data Scientist.

In Crystal we believe that to find such a Data Scientist Candidate is almost impossible, as a set of knowledge and skills for the ideal candidate is incompatible. Thus, market attempts to search and find the ideal Data Scientist could be described in some way as Mission: Hunting a Data Scientist = Hunting a Unicorn?

Figure 8 The Ideal Data Scientist



That is why, at Crystal, we developed our approach how to build and select the Data Scientist Talents.

### CRYSTAL PRAGMATIC APPROACH TO BUILD AND SELECT DATA SCIENTIST TALENTS

**CRYSTAL PRAGMATIC APPROACH TO SELECT DATA SCIENTIST TALENTS** - is a set of pragmatic recommendations for business how to attract a right DS candidate.

Crystal Pragmatic Approach to select Data Scientist Talents is based on the following principles:

**1) DATA SCIENCE PROJECTS** as multidisciplinary projects looking at specific business issue/problems (please note: we start with business problem not with tech)

**2) DATA SCIENCE PROJECTS ARE HIGHLY COLLABORATIVE WITH A DIVERSE EXPERT TEAM**, including the following disciplines and key roles:

- **BUSINESS EXPERTISE** - “The Business Guy”, it may be an accountant, financial analyst, engineer, marketing manager or other business expert who has a solid understanding of the business processes and preferably with the academic background in the business domain, and who has a system view of the business problem;

- **MATHEMATICS & STATISTICS** - “The Data Nerd”, a professional with profound knowledge and good skills in building, implementing and developing predictive and prescriptive models, and with a strong fundamental understanding of various modern machine-learning methods;

- **COMPUTER SCIENCE / IT** - “The Hacker”, it may be a Data Developer/Software Engineer or other professional with experience in Big Data, ML technologies, general purpose programming languages, and who is a tech solver of the business issue.

- **COMMUNICATION** - “The Communicator”, it may be a Data Analyst or a business expert with

strong oral and written communication skills, and who is able to communicate sophisticated concepts to intelligent business community.

3) Adopting a new approach from searching the candidate to **SELECT THE DATA SCIENTIST TALENT BY PLACING THE PRAGMATIC JOB OFFER**. Once the employer defines the Data Science Project with the business problem at its core (principle 1), he can identify the team players needed with the targeted Data Science candidate and thus place the Pragmatic Job Offer that is highly personalized and attractive for the right Talent.

By using the Crystal Pragmatic Approach to select Data Scientist Talents, companies will be able to shift the paradigm from searching a Unicorn to selecting and attracting a targeted Data Scientist Talent, as they will start considering the process of “doing Data science” as an interdisciplinary project and thus recruit a right candidate for the Data Science team.

**CRYSTAL PRAGMATIC APPROACH TO BUILD DATA SCIENTIST TALENTS** – is an educational methodology based on more than 15 years’ experience to build Talents through attracting and

selecting them from the best public universities, then educating Talents in Crystal Academy via professional courses and project trainings, followed by further service delivering.

Focusing on a market problem-solving and basing on its educational expertise, Crystal System is planning to implement the Data Scientist course (DSC) with its partner universities, made in collaboration with the best domain experts, aligned with the market business requirements and Crystal System partners, and it is supported by the market research.

The Crystal System DSC is aimed to build and accelerate Talents career in Data Science providing them with the world-class training and skills required to become successful in this domain, as well as prepare the industry-ready professionals and teams.

The Crystal DSC is focused on both market and Talents: by deep diving Talents into the nuances of data interpretation, mastering powerful programming skills and focusing on the business acumen as an essential element, and helping market by building industry-ready professionals (Figure 9).

(Figure 9) Crystal System Pragmatic Approach to Build Data Scientist Talents

