dataVA Datathon 2016  
Question Analysis Report

Bernard Schmidt, Jim Dugan, Ashfaq Khan, & Varaha Santosh Malla Naga

# Overview

On August 2, 2016, the CCARS project team met to discuss the eight “burning questions” highlighted on the datathon registration home page. The analysis below is the end result of the team discussion. The question discussion took the entire 2 ½ hour period, leaving no time to discuss the six “challenges to consider” also posted on the datathon registration home page.

## Question 1

What are the skills required of jobs? How do those skills align to available training and curricula at Virginia’s community colleges?

### Data Required

#### VT Jobs Data

Issue – The skills column is sparsely populated. The job description contains the skills, but can it be semantically parsed to reliably extract out skills for each of the jobs?

Issue – Is there a list of standard skills (e.g., is JAVA a skill? Are we assuming technical skills only - what about soft skills?)? What about certifications (e.g., M.D., R.N., CISSP, PMP, etc.)?

#### VCCS Training and Curricula Data

Issue – The VCCS website does not contain data with sufficient information. The VCCS website only has available courses and abstract descriptions – no detailed course content summaries.

Issue – The VCCS website has no centralized Workforce Development training information. It instead contains links to the workforce development web pages of the 23 individual community colleges in the system.

### What is *really* being asked and what is *really* being solved here?

This is really a policy question for the VCCS Chancellor regarding curricula and training being offered by the VCCS.

From the perspective of someone without a college degree asking this question, will this information give them an understanding of which Community College degree or training program which will lead to an employable job?

### Is it Solvable?

Assuming that skills are listed in the job rec or description – and those skills are also listed in the VCCS curricula/training data – then we can weigh the skills of the jobs vs. the curricula.

What about soft skills (e.g., verbal and written communications) – VCCS course abstracts and course content summaries do not indicate whether a particular curricula or training course will give students those particular skills. How can we, or even can we, infer those skills for any particular curricula or training course?

Do “standard” skill definitions exist that can be applied to solve this problem (e.g., what does JAVA mean? What does NOSQL mean? Etc.) Can we create them?

## Question 2

What jobs require an academic certificate or associate degree in addition to a license or industry certification? What jobs require a high school diploma and a license or industry certification?

### Data Required

#### VT Jobs Data

Issue – The skills column is sparsely populated. The job description contains the license and industry certification information, but can it be semantically parsed to reliably extract out the license and industry certification for each of the jobs?

Issue – Is there a list of standard licenses and industry certifications?

### What is *really* being asked and what is *really* being solved here?

This is really a question for someone graduating from high school or community college (e.g., Manassas High School Automotive Program or NVCC Automotive Program) looking for a job they can qualify for with a particular degree – especially outside of Northern Virginia (e.g., southwestern Virginia).

### Is it Solvable?

Yes, assuming we have a standardized list of licenses and industry certifications. Where can that list be found (e.g., licensed healthcare provider, licensed daycare provider, licensed DPOR contractor, morticians, ASE certified mechanics, etc.)?

## Question 3

What is the demand for degrees and credentials by level and type? How are they connected to education programs and majors?

### Data Required

#### VT Jobs Data

Issue – The skills column is sparsely populated. The job description contains the skills, but can it be semantically parsed to reliably extract out skills for each of the jobs?

Issue – Is there a list of standard skills (e.g., is JAVA a skill? Are we assuming technical skills only - what about soft skills?)? What about certifications (e.g., M.D., R.N., CISSP, PMP, etc.)?

#### Academic Education Programs and Majors

Issue – Access to VCCS, non-profit and for-profit University catalog degree programs would be required. The information is available from the State Council of Higher Education for Virginia (SCHEV) via a web form, but can SCHEV provide it in an easy-to-process format?

### What is *really* being asked and what is *really* being solved here?

This is really a policy question for academic administrators (e.g., VCCS Chancellor, University President, Academic Dean of a University College, and SCHEV).

### Is it Solvable?

Yes, assuming we have the right data elements.

## Question 4

What are the real workforce needs of employers? Are those needs adequately reflected in job postings?

### Data Required

#### Employer Workforce Needs

Issue – The data does not exist.

### What is *really* being asked and what is *really* being solved here?

This is really a business organization question that can only be addressed internally within a business.

### Is it Solvable?

No.

## Question 5

What are the historical patterns for jobs in terms of skill requirements, educational attainment, and credentials? (basis for predictive analytics)

### Data Required

#### Multi-year VT Jobs Data

Issue – The skills column is sparsely populated. The job description contains the skills, but can it be semantically parsed to reliably extract out skills for each of the jobs?

### What is *really* being asked and what is *really* being solved here?

This is really a question for academic administrators wanting to know whether their academic and training programs are up-to-date or ahead-of-the-curve.

### Is it Solvable?

On one hand, yes, assuming we can extract historical data from the VT Jobs Data, we can map generic job categories (e.g., programmer, because specific programmer jobs titles will evolve over time) with changing skills. We can show trends, but not predictions.

On the other hand, no, predictions require information not contained in the jobs data – requires information about future jobs which does not currently exist or requires complex analysis of technology trends.

## Question 6

How do we move the conversation and measurement on employability beyond wages? What are other metrics and how/why are they actionable?

### Data Required

Unknown.

### What is *really* being asked and what is *really* being solved here?

The question can be viewed either from the perspective of the employer and/or the employee. The question, as posed, is too vague to determine how best to answer it.

### Is it Solvable?

No.

## Question 7

For all the noise on soft skills how do we validate that they are teachable rather than learned by experience?

### Data Required

Unknown.

### What is *really* being asked and what is *really* being solved here?

The question can be viewed either from the perspective of the employer and/or the employee. The question, as posed, is too vague to determine how best to answer it.

### Is it Solvable?

No.

## Question 8

What is the best way to get training and “skill up” into a better paying position?

### Data Required

Tied to Question 5 trends.

### What is *really* being asked and what is *really* being solved here?

The question can be considered from an employee attempting to improve their employment situation. However, there are a number of traditional (e.g., academic degrees, certificates, and workforce training) and non-traditional (e.g., MOOCs such as edX, Coursera, Udemy, Kahn Academy, etc.) approaches to consider when considering this question.

### Is it Solvable?

The question is not solvable with the available data; but, also because it depends on specific employer requirements for acceptable training or “skill up.” Are non-traditional approaches acceptable to employers?