

Quismorio, Brenda A.; Pasquin, Maria Antonette D.; Tayco, Claire S.

Working Paper

Assessing the alignment of Philippine higher education with the emerging demands for data science and analytics workforce

PIDS Discussion Paper Series, No. 2019-34

Provided in Cooperation with:

Philippine Institute for Development Studies (PIDS), Philippines

Suggested Citation: Quismorio, Brenda A.; Pasquin, Maria Antonette D.; Tayco, Claire S. (2019) : Assessing the alignment of Philippine higher education with the emerging demands for data science and analytics workforce, PIDS Discussion Paper Series, No. 2019-34, Philippine Institute for Development Studies (PIDS), Quezon City

This Version is available at:

<https://hdl.handle.net/10419/240985>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

DISCUSSION PAPER SERIES NO. 2019-34

Assessing the Alignment of Philippine Higher Education with the Emerging Demands for Data Science and Analytics Workforce

*Brenda A. Quismorio, Maria Antonette D. Pasquin,
and Claire S. Tayco*



The PIDS Discussion Paper Series constitutes studies that are preliminary and subject to further revisions. They are being circulated in a limited number of copies only for purposes of soliciting comments and suggestions for further refinements. The studies under the Series are unedited and unreviewed. The views and opinions expressed are those of the author(s) and do not necessarily reflect those of the Institute. Not for quotation without permission from the author(s) and the Institute.

CONTACT US:

RESEARCH INFORMATION DEPARTMENT
Philippine Institute for Development Studies

18th Floor, Three Cyberpod Centris - North Tower
EDSA corner Quezon Avenue, Quezon City, Philippines

publications@mail.pids.gov.ph
(+632) 8877-4000

<https://www.pids.gov.ph>

Assessing the Alignment of Philippine Higher Education
with the Emerging Demands for Data Science
and Analytics Workforce

Brenda A. Quismorio
Maria Antonette D. Pasquin
Claire S. Tayco

PHILIPPINE INSTITUTE FOR DEVELOPMENT STUDIES

December 2019

Abstract

Rapid advancement in technology has allowed for far-reaching use of data. This has consequently led to an increasing demand for Data Science and Analytics (DSA) professionals. However, recent studies show that such demand is often not met in many economies. Such DSA skills shortage is claimed to be rooted in the mismatch between the skills the industry demands and the skills academic institutions supply. This mismatch is evident in the Philippines where studies also reveal certain difficulties of current Philippine education and training to meet the level of competencies required to do high-skilled jobs. An indicator of this weak point is the persistent high youth unemployment and underemployment rate in the Philippines where graduates land jobs which their completed education did not intend for them.

As the first step in addressing this shortage of industry-ready DSA workers, it is necessary to know the DSA skills demanded by the industry and the DSA skills with which academic institutions equip their students. To do this, the study employed the Analytics Association of the Philippines' (AAP) Professional Maturity Model, which is based on the ten APEC-recommended DSA competencies, as analytical framework. It estimated the current availability of DSA competencies using available data from the Labor Force Survey (LFS) and interviews with companies engaged in analytics activities. Meanwhile, the profiling into four DSA job roles (Data Steward, Data Engineer, Data Scientist and Functional Analyst) of these workers was done with the use of relevant data from online job postings. On the other hand, the current supply for DSA job roles were mainly determined from survey interviews of analytics practitioners and undergraduate program administrators and CHED's databases.

Findings showed that in terms of the demand-side, the Functional Analyst job role emerges as the most sought role of employers among the four DSA-related job roles. The demand for these DSA-roles also mostly come from the Information and Communication industry group under the Services sector. For each demanded job role, the employers' require differing DSA competencies and levels of proficiency. The findings also revealed that one challenge for employers in hiring DSA talents is the inadequate 21st century skills more than other technical skills.

With regard to the supply-side, top ten undergraduate programs were identified to be DSA – related programs based on the common degrees of current analytics practitioners. The extent to which these DSA-related programs prepare their respective graduates with the basic proficiency of each competency varies. However, the curricula assessments reveal that these degree programs mostly fully equip their graduates with data engineering and statistical techniques competencies and correspondingly fully enable them to perform the Data Engineer job role. Out of all DSA-enabled graduates from these DSA-related programs, some graduates were also assessed as industry-ready while others still need training for some DSA competencies.

With these results, the study pointed to a misalignment between the demand and supply of the DSA workforce in the country. Specifically, there are DSA competencies – both sought by employers and required by the analytical framework of the study – that are poorly supplied by these DSA-related undergraduate programs. Moreover, while employers mostly look for Functional Analysts, most of the graduates of the degree programs are fully enabled to perform the Data Engineer job role. Because of these findings, the study offers two main recommendations. First, it recommends the use of the AAP Professional Maturity Model to

define the DSA profession. To allow the growth and maturity of the still infant analytics industry, there is a need to address the current lack of a common definition of the analytics profession among stakeholders. Having a common understanding can facilitate efforts to align supply and demand and eventually the maturity of the industry. With the first recommendation as starting point, the study further endorses the promotion of government-industry-academe linkages to expand the existing market for DSA workforce in the country.

Keywords: data science and analytics, analytics, digital transformation, skills mismatch

Table of Contents

1. Introduction	1
2. Objective/s	2
3. Significance of the Study	3
4. Scope and Limitation	3
5. Review of Related Literature	3
5.1. Digital transformation, big data and skills mismatch in the Philippines	3
5.2. Data Science and Analytics (DSA)	5
6. Definition of terms	8
7. Research design and methodology	8
7.1. Conceptual framework	9
7.2. Research Design, Instruments, and Participants	10
7.2.1. Demand-side	10
7.2.2. Supply-side	12
8. Results and Discussion	13
8.1. <i>Demand-side</i>	13
8.1.1. Current availability of DSA competencies	13
8.1.2. Current demand DSA job roles	15
8.1.3. Work analysis of in-demand DSA job roles	16
8. 2. Supply side	18
8.2.1. DSA-related undergraduate degree programs	18
8.2.2. Current supply of DSA competencies and DSA-job ready graduates	20
8.2.2.1. Bachelor of Science in Computer Science	20
8.2.2.2. Bachelor of Science in Business Administration (BS BA)	21
8.2.2.3. Bachelor of Science in Statistics	22
8.2.2.4. Bachelor of Science in Math/Applied Math	23
8.2.2.5. Bachelor of Science in Information Technology (BS IT)	23
8.2.2.6. Bachelor of Library and Information Science (BLIS)	24
8.2.2.7. Bachelor of Science in Physics/Applied Physics	25
8.2.2.8. Bachelor of Science/Bachelor of Arts in Economics	26
8.2.2.9. Bachelor of Science in Industrial Engineering (BSIE)	27
8.2.2.10. Bachelor of Science in Civil Engineering (BSCE)	27
8.2.3. DSA-related Degree Programs and the 21 st Century Skills	28
8.2.4. Summary of Supply-side discussion	31
9. Conclusion: Demand and Supply	33
10. Recommendations	34

List of Tables

Table 1. Projected DSA workforce demand in select economies	4
Table 2. Data Science and Analytics job roles	6
Table 3. The Analytics Association of the Philippines' (AAP's) professional maturity model	10
Table 4. Employment size of industry Subclasses with mostly analytics-related activities	14
Table 5. Job profile of in-demand job roles	16
Table 6. Work Matrix of DSA Job Roles	18
Table 7. List of Data Science and Analytics-related undergraduate degree programs	19
Table 8. In-demand Occupations by Minimum Requirement	19
Table 9. BS Computer Science DSA competencies and job roles	21
Table 10. BS Business Administration DSA competencies and job roles	22

Table 11. BS Statistics DSA competencies and job roles	22
Table 12. BS Math/Applied Math DSA competencies and job roles	23
Table 13. BS Information Technology DSA competencies and job roles	24
Table 14. BS Library and Information Science DSA competencies and job roles	25
Table 15. BS Physics/Applied Physics DSA competencies and job roles	25
Table 16. BS Economics DSA competencies and job roles	26
Table 17. BS Industrial Engineering DSA competencies and job roles	27
Table 18. BS Civil Engineering DSA competencies and job roles	28
Table 19. 21 st century skills component of DSA-related undergraduate programs' actual curricula	29
Table 20. 21 st century skills component of DSA-related undergraduate programs' CMO curricula	30
Table 21. Results of 2018 Philippine Talent Map Initiative	31
Table 22. Graduates equipped with DSA competencies	32
Table 23. Job-ready graduates from DSA-related undergraduate programs	33
Table 24. Distribution for Demand and Supply of DSA Job Roles	33

List of Figures

Figure 1. Conceptual framework: the alignment between demand and supply of DSA competencies in the Philippines	8
Figure 2. Skills Extraction and Job Segmentation Process	11

References

37

Appendices

40

Appendix A. Detailed Online Job Scraping methodology	40
Appendix B. List of pre-determined set of skills by Wang (2017)	50
Appendix C. Distribution table of DSA employees' degree programs	50
Appendix D. Employment Share per Major Industry Group (2013-2017)	51
Appendix E. Employment Size of Industry subclasses of BPM subsector	52
Appendix F. Employment size of Analytics-related activities from PSIC's list and definition of industries (2013-2016)	54

Assessing the alignment of Philippine higher education with the emerging demands for data science and analytics workforce

*Brenda A. Quismorio, Maria Antonette D. Pasquin,
and Claire S. Tayco**

1. Introduction

The rapid and continuous advancement in technology allowed big data¹ to be more accessible nowadays for decision-makers in governments, businesses or any organization. McKinsey notes that “data [have] become widespread across industries that [they] can now be considered as an important factor of production [along] with labor and capital (McKinsey as cited in Tableau 2016, p. 2).” Because of this explosion of data, the field of Data Science and Analytics (DSA) has also gained much relevance and attention. “Companies of all sizes rely on data science and analytics as key transformational components to their core operations (Business Higher Education Forum 2014 p. 1).” The recent study of Microsoft about Digital transformation in the Asia Pacific shows that big data and analytics topped as the third mostly invested technology in 2018 next to Security and Mobility (Microsoft, 2018).

This growing interest in DSA consequently causes an uptrend in demand for professionals from this field. However, recent studies reveal that such demand is often not met. Developing a workforce with DSA skills continue to be a challenge for many economies (Asia Pacific Economic Cooperation Human Resource Development Working Group, 2017; Pricewaterhouse Cooper and Business-Higher Education Forum, 2017). In fact, next to Cybersecurity skills, DSA skills are identified by 95% of the employers as most difficult to find among prospective employees in the United States (PwC & BHEF, 2017).

The DSA skills shortage can be traced to the mismatch between the skills that the industry demands and the skills that academic institutions supply. Most higher education data science programs are lodged under the school of engineering and/or the department of computer science (BHEF, 2014). What makes this practice quite problematic is that it may prepare data science experts who while possessing the quantitative skills, still often lack the domain-specific knowledge in health, transportation, economics, business, and public policy that companies need (BHEF, 2014). At the same time, this may contribute to having a workforce who may only know the business domain but lack the technical skills. Hence, employers usually encounter difficulties looking for the right DSA talents, thus, perpetuating the current problem of DSA skills shortage.

As this skills mismatch may lead to economic losses, initiatives have been started to equip the future workforce with the skills of data science and analytics (APEC Human Resources Development Working Group, 2019). For instance, the United States has started to gear up its universities with as many DSA degree programs (Tableau, 2016). The APEC Human Resource

* Assistant Professor (University of Asia and the Pacific), Junior Researcher (University of Asia and the Pacific), and Senior Analytics Researcher (Cirrolityx), respectively.

¹ As cited in Dadios, et. al (2018), a scoping study on the Fourth Industrial Revolution, there is yet no universal consensus on the definition of “big data.” Nonetheless, it is understood by most players in the industry as “the digital data sets that are so large (either in terms of scale, i.e. “volume”, or streams across time, i.e. “velocity”) or complex (in terms of variety)” which are mostly sourced from use of digital gadgets or social media, etc.

Development Working Group (2017) recommended a set of DSA competencies to serve as a reference for academe, the industry and government sectors in institutionalizing DSA.

With the advent of increasing demand for Data Science and Analytics, the lack of employment opportunities, often claimed for graduates of tertiary education in the Philippines, can be addressed. However, this idea still posts difficulties since DSA skills are still poor or underdeveloped among the workforce in the Philippines (Tan & Tang, 2016). In addition, the level of competencies of high-skilled jobs required by industries is not met by current Philippine education and training. Specifically, the Philippines suffers from “outdated curriculum, weakness in STEM education and lack of soft skills training (including proficiency in English), insufficient collaboration between the educational institutions and industry, inadequate instructional facilities, poor quality of instructors who often do not possess the necessary knowledge or industry experience to impart industry-relevant skills to the students, lack of industry exposure for students, and the absence of a rigorous skill competency certification program” (Tan & Tang 2016, p. 100).

While existing studies recognize the data science and analytics skills mismatch in the Philippines, there continues to be a dearth in assessment studies about the current state of data science and analytics competencies in the country. This, in return, makes policy formulation challenging across relevant stakeholders such as business, government, and educational institutions. This paper thereby addresses such literature gap by examining how the Philippines fares in equipping the workforce for future jobs. Specifically, this paper, provides an introspection of the Philippine education curricula on DSA-related degree programs insofar as they equip graduates with the right set of competencies needed in their future workplace.

2. Objective/s

This paper aims to assess existing Philippine education specifically the curricula of Data Science and Analytics-related undergraduate degree programs insofar as they meet the DSA competency requirements of the industry. In doing so, the paper augments the main question with the following specific questions:

Are Philippine undergraduate programs able to supply the current demands for Data Science and Analytics (DSA) talents?

i. What is the current demand for DSA-related jobs in the Philippines?

- a. What DSA job roles do companies look for?
- b. What set of competencies do employers require from DSA workers?
- c. What challenges do companies face when hiring these jobs?
- d. What is the nature of these jobs in accordance with employers' requirements and framework's standard.

ii. What is the current supply of DSA-related jobs in the Philippines?

- a. What are the DSA-related undergraduate programs?
- b. Which DSA jobs do these programs prepare their graduates to perform?

- c. To what extent do these programs equip their students with the basic proficiency of the DSA competencies upon graduation?

3. Significance of the study

How can the Philippines prepare its workforce for the “Fourth Industrial Revolution²” or future jobs? This study takes the first step in assessing the current ability of Philippine educational institutions in enabling the future workforce for the Data Science and Analytics (DSA) jobs needed by industries. This study then aims to recommend courses of actions that education, industry and government sectors can do to ensure the availability of adequate supply of DSA workforce. In addition, this paper sets the stage for further studies on DSA in the Philippines given the dearth in the literature on this subject.

4. Scope and limitation

Ten DSA-related undergraduate programs were identified from the ten top undergraduate programs of surveyed DSA practitioners. Each DSA-related undergraduate program was evaluated by academe experts using two of its curricula: the most recent actual curriculum of their university and the sample curricula proposed by the Commission on Higher Education (CHED). While the sample curriculum contains the minimum requirements of the degree program and serves as a guide for Higher Education Institutions (HEIs) in developing their actual curriculum, the evaluation of the sample curriculum does not consider other factors that affect the delivery of instruction such as teacher qualifications, teaching materials, facilities and evaluation tools, etc.

Data on the demand for DSA workforce is very scarce given that the field is just beginning. The data used were mostly sourced from government agencies’ published reports on employment statistics and employers’ online job postings which points to the demand coming mostly from the Information Technology and Business Processing Management (IT-BPM) industry. Other industries, though they may also demand for DSA workforce, are not included in the study.

5. Review of related literature

5.1 Digital transformation, big data and skills mismatch in the Philippines

The wave of digital transformation crashes across various economies and the Philippines is not exempt from such phenomenon. In fact, the recent study of Microsoft (2018) highlights the economic impact that digital transformation may bring to the country. It projects that by 2021; about 40% of the Philippine Gross Domestic Product (GDP) would be credited to digital transformation. Specifically, it would contribute 0.4% of the projected 6.9% growth in GDP. The same study nonetheless enumerates other positive socio-economic consequences such as increase in income and higher-value jobs in the future, and a rise in opportunities for individuals through better access to education and training. Furthermore, Microsoft (2018) presents five potential benefits of digital transformation for the companies of Southeast Asian countries,

² Refer to Dadios et. al (2018), PIDS’ Scoping Study on the Fourth Industrial Revolution for further explanation of the phenomenon.

namely: improved profit margin; increased revenue from new products and services; improved customer advocacy; loyalty and retention; improved productivity; and reduced costs.

Microsoft (2018) mentions that rather than displacement, job transformation would take place. In the next three years, 92% of the Philippine jobs would be transformed. Of these transformed jobs, 34% would be retained and upskilled, 31% would be new roles due to digital transformation investment and 27% would be outsourced, automated or made redundant. Meanwhile, only 8% of the jobs would remain unchanged. The sectors most vulnerable to automation, according to Francisco et. al (2019) are 1.) youth (age 15-24), 2.) those who have received less education, 3.) those with lower incomes, 4.) those employed in casual or irregular and seasonal jobs or working without pay and 5.) those belonging to the agriculture, forestry and fishing sectors.

With digital transformation comes the exponential growth of data. BHEF (2014, p. 1) quotes Gartner Inc.'s projection that "in less than 12 months, 4.4 million information technology (IT) jobs to support big data will be created globally. About 1.9 million of those jobs will be within the United States, and big data has the potential to create three times that number of jobs outside of IT." In 2017, the growth of big data in the United States was projected at 40% each year with potential value of \$300 billion to the nation's health care industry alone (BHEF, 2017).

This increasing demand for DSA workforce was likewise observed in the APEC region (APEC Human Resource Development Working Group, 2017; BHEF, 2014; BHEF, 2017). At the top of skills shortage among the APEC economies are Data Science and Analytics (DSA) skills (APEC Human Resource Development Working Group, 2017). The current and projected demand for DSA-related workers in select APEC economies are presented in Table 1.

Table 1. Projected DSA workforce demand in select economies

Economy	Current DSA Workers	Projected DSA Workers Needed	Percent Change
Malaysia	4, 000 (2016)	20, 000 (2020)	400%
Philippines	147, 420 (2016)	340, 880 (2022)	131%
Singapore	9, 300 (2015)	15, 000 (2018)	61%
Canada	33, 600 (2016)	43, 300 (2020)	33%
United States	2, 350, 000 (2015)	2, 720, 000 (2020)	16%

Source: APEC Human Resource Development Working Group (2017)

The skills mismatch between the demand and the supply of needed skills (skill shortage or gap) that gives rise to difficulties in recruitment processes (Iredale et. al as cited in Ramos 2016) signifies the inability of educational institutions to meet industry demands (Tan & Tang, 2016; Elemia, 2016). This “school-industry gap,” an arguably most serious hurdle in developing an effective workforce, is felt across industries (Tan & Tang 2016, p. 108).

The Labor and Employment Statistics Survey of the Philippines (PSA, 2012/2013) especially highlights the shortage of high skills, as those in the field of Science and Technology (Ramos, 2016). Specifically, graduates who join the Information Technology-Business Processing Outsourcing (IT-BPO) industry and emerging infrastructure management services lack the necessary IT skills and Infrared (IR) software proficiency (Tan & Tang, 2016). Moreover, skills gaps that include “numerical competence, verbal and report writing skills, familiarity with different business models and terms, industry- specific knowledge and processes codes and terms” also impair the growing non-voice segments such as Knowledge Processing Outsourcing (KPO) of business analytics, insurance services, and health information management (Tan& Tang, p. 95).

Earlier efforts and interventions³ have been done throughout the years. Recent efforts to augment the gap were manifested in the proliferation of Data Science degree programs in various universities and online courses on Data Science such as Coursera, a MOOC (Massive Open Online Course) and Microsoft’s Professional Degree in Data Science (Tan & Tan, n.d.). Hence, it is further suggested that this rising demand be addressed by encouraging schools with existing B.S. Statistics and B.S. Computer Science degrees to also consider offering a B.S. Data Science (Tan & Tan (n.d.).

Despite these initiatives, the problem persists as reflected in the rising youth unemployment. Such situation, thus, warrants the assessment of current Philippine education on its ability to equip its graduates with the set of competencies needed by industry sectors. This paper, thereby, seeks to address this gap by assessing, in particular, the DSA -related degree programs currently offered by various Higher Education Institutions to find out whether or not there is indeed skills shortage on DSA competencies.

5.2. Data Science and Analytics

Data Science and Analytics (DSA), as a new field, has no universally accepted definition. One reason provided for this is that DSA involves a wide-range of skills or competencies that could not easily be identified or narrowed down to a single definition (Demchenko, 2017). Nonetheless, existing literature on DSA provides a host of definitions.

Metamarket CEO Mike Driscoll defines Data Science as “the civil engineering of data (Schutt & O’Neil 2013, p. 7). He further claims that definitions following this idea of a Data Science, such as Conway’s Venn Diagram of Data Science (2010), possess both the practical knowledge of tools and materials and theoretical understanding of the multidisciplinary nature of the field (Schutt & O’Neil 2013).

³ In 2014, it can be recalled how the Aquino administration, through the Department of Labor and Employment, made its Thrust and Priorities centered on addressing the job-skill mismatch in the country. The Human Development and Poverty Reduction Cluster namely, education and curriculum review; development of a Philippine Qualifications Framework (PQF); implementation of career guidance advocacy; and optimizing the utilization of the Enhanced Phil-Job.Net (E-PJN) were the four identified programs under such policy action. Earlier on, in 2013, the implementation of the K-12 Program was also a major step of the government to address the problem of skill mismatches in the Philippines. More recently, another Senate Bill 1456 or the Philippine Qualifications Framework (PQF) Act of 2017, which just passed in its 3rd and final reading also served as a policy action towards job-skill alignment (Pasion, 2017).

Meanwhile, the National Institute of Standards and Technology (NIST) distinguishes the meaning of Data Science and Analytics. “Analytics” is defined as a step in the data life cycle that involves the collection of raw data, preparation of information then analytics, visualization and access (“synthesis of knowledge from information”). On the other hand, “Data Science” refers to the “extraction of actionable knowledge directly from data through either a process of discovery, or hypothesis formulation and hypothesis testing (NIST as cited in PwC and BHEF 2017, p. 3).”

Using these two concepts “analytics” and “data science, PwC and BHEF (2017) formulate the landscape of DSA job roles as seen in Table 2 below”. It classifies DSA job roles into analytics-enabled jobs and data science jobs. Professionals who perform analytics-enabled jobs are Data-driven Decision Makers who “leverage data to inform strategic and operational decisions” and Functional Analysts who “utilize data and analytical models to inform domain-specific functions and business decisions (PwC and BHEF 2017, p. 4).” On the other hand, professionals who perform data science jobs are the Data Engineers who “design, build, and maintain an organization’s data and analytical infrastructure,” the Data Analysts who “leverage data analysis and modeling techniques to solve problems and glean insight across functional domains” and the Data Scientists who “create sophisticated analytical models used to build new data sets and derive new insights from data (PwC and BHEF 2017, p. 4).” Job titles currently used in organizations were mapped to these five Data Science and Analytics Job roles.

Table 2. Data Science and Analytics job roles

Data Science and Analytics Job Roles				
Analytics-enabled Jobs		Data Science Jobs		
Data-driven decision-makers	Functional Analysts	Data Engineer	Data Analyst	Data Scientist
Common job titles: Chief Executive Officer Chief Data Officer Chief Information Officer Director of IT Financial manager Human resources manager Marketing manager	Common job titles: Actuary Business/Management analyst Compensation/Benefits analyst Financial analyst Geographer/GIS specialist HRIS analyst Operations analyst Researcher	Common job titles: Business intelligence architect Computer systems engineer Data warehousing specialist Data administrator Database architect Systems analyst	Common job titles: Data mining analyst Business intelligence analyst	Common job titles: Biostatistician Data engineer Data scientist Financial quantitative analyst Statistician

Source: PwC and BHEF (2017, p. 4)

Demchenko (2017) extends the discussion and provides a more detailed and comprehensive definition of the Data Science Profession Profiles encapsulated in the EDISON Data Science Framework (EDSF). EDSF presents five core Data Science competencies and skills groups as follows:

- Data Science Analytics (including Statistical Analysis, Machine Learning, Data Mining, Business Analytics, others)
- Data Science Engineering (including Software and Applications Engineering, Data Warehousing, Big Data Infrastructure and Tools)
- Domain Knowledge and Expertise (Subject/Scientific domain related)
- Data Management and Governance (including data stewardship, curation, and preservation)
- Research Methods for research related professions and Business Process Management for business related professions

The APEC Human Resource Development Working Group (2017, p. 3) adopts the definition of Data Science and Analytics as “the ability to gather, analyze, and draw practical conclusions from data, as well as communicate data findings to others.” Likewise, the definitions of three basic concepts in DSA, namely Data Science, Big Data and Data Analytics were given as (APEC Human Resource Development Working Group, 2017, p. 5):

- Data Science: The data science field addresses both structured and unstructured data in terms of data cleansing, preparation, and analysis. Overall, data science is a grouping of techniques that enable insight and information extraction from data. Internet searches, search recommenders, and digital advertisements and profiles all fall under the application of data science.
- Big Data: The big data field addresses the massive volumes of data that are generated across industries; the majority of this data is simply too large to be processed or interpreted effectively with traditional approaches. The analysis of big data can deliver insights to support better decision-making and business strategies. Financial services, retail, and communication all generate big data and thus require skilled techniques.
- Data Analytics: The data analytics field examines raw data to draw conclusions from the information. Inference skills are necessary within data analytics, as the researcher must derive conclusions from what he or she already knows.

Two definitions in the Philippine literature come from Tan & Tan (n.d, p.4) and the Analytics Association of the Philippines. For Tan & Tan (n.d, p.4), DSA is “a systematic study of digital data using statistical techniques and applications of computer science. Its goal is to make sense of vast amounts of dynamic data and extract information that will lead to new knowledge that can provide actionable insights for decision makers.” On the other hand, the recently established Analytics Association of the Philippines (AAP) provides the operational definition of “Analytics” as “a process of progressing data along the value chain as it transforms data to information to insight to imperatives (or actionable insights) with the purpose of delivering the right decision support to the right people and digital processes at the right time for the good (AAP 2018).”

6. Definition of terms

In the absence of an agreed definition, this study takes on the operational definition of “Analytics” by AAP for Data Science and Analytics (DSA). It is deemed that with this definition, the specific attributes identified by varying references about DSA are captured. In this paper, the discourse about the field of DSA is further explored especially with regard to the set of competencies supplied by HEIs and demanded by companies in the Philippines. It is

through this approach that the paper attempts to contribute to the on-going development of the field of DSA and narrowing the gap between what DSA is in theory and practice.

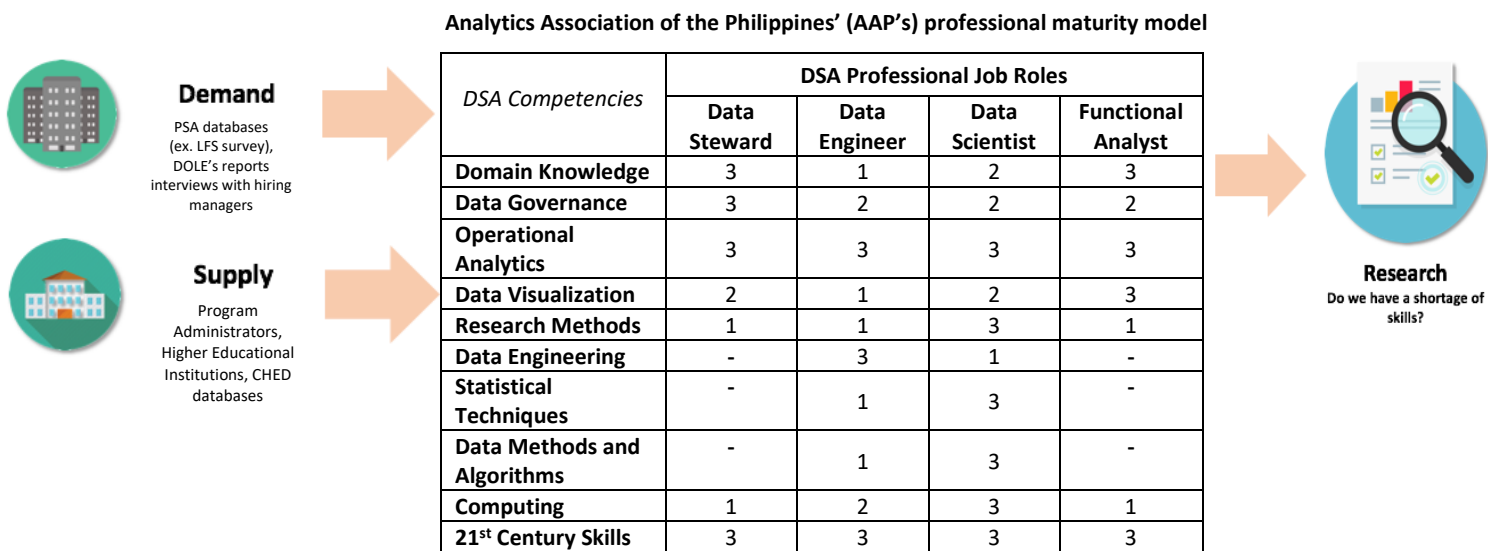
Analytics is “a process of progressing data along the value chain as it transforms data to information to insight to imperatives (or actionable insights) with the purpose of delivering the right decision support to the right people and digital processes at the right time for the good (AAP 2018).”

7. Research design and methodology

7.1. Conceptual framework

The conceptual framework (Figure 1) makes use of the Professional Maturity Model of the Analytics Association of the Philippines (Table 3) as the point of alignment between the supply and demand for the DSA workforce.

Figure 1. Conceptual framework: the alignment between demand and supply of DSA competencies in the Philippines



Source: AAP (2018); Authors' illustration

The model defines four DSA job roles, namely Data Steward, Data Engineer, Data Scientist and Functional Analyst (Table 3).

- **Data Steward:** Develops, enforces, and maintains an organization's data governance process to ensure that data assets provide the organization with high-quality data.
- **Data Engineer:** Designs, constructs, tests, and maintains data infrastructures including applications that extract, transform, and load data from transactional systems to centralized data repositories.
- **Data Scientist:** Leverages statistical techniques and creates analytical models to derive new insights from quantitative and qualitative data.
- **Functional Analyst:** Utilizes data and leverages on derived insights to help organizations make better decisions on a specific functional domain.

Each job role is characterized by the required proficiency level of the 10 APEC Recommended DSA competencies. These set of DSA competencies were developed by a 50-person Advisory Group composed of representatives from the industry, academe, and governments of 14 various APEC-member countries (APEC Human Resource Development Working Group 2017).

- Operational Analytics - use general and specialized business analytics/intelligence techniques for the investigation of all relevant data to derive insight for decision-making.
- Data Visualization and Presentation - create and communicate compelling and actionable insights from data using visualization and presentation tools and technologies.
- Data Management and Governance - develop and implement data management strategies, incorporating privacy and data security, policies and regulations, and ethical considerations.
- Domain Knowledge and Application - apply domain-related knowledge and insights to effectively contextualize data, achieved by practical experience and exposure to emerging innovations.
- Statistical Techniques - apply statistical concepts and methodologies to data analysis.
- Computing - apply information technology and computational thinking, and utilize programming languages and software and hardware solutions for data analysis.
- Data Analytics Methods and Algorithms - Implement and evaluate machine learning methods and algorithms on the data to derive insights for decision-making.
- Research Methods - utilize the scientific and engineering methods to discover and create new knowledge and insights.
- Data Science Engineering Principles - use software and system engineering principles and modern computer technologies to develop data analytics applications.
- 21st Century Skills - exhibit cross-cutting skills essential for Analytics at all levels, including but not limited to: collaboration, ethical mindset, empathy, social and societal awareness, dynamic (self) re-skilling, and entrepreneurship.

Table 3. Analytics Association of the Philippines' (AAP's) professional maturity model

DSA Competencies	DSA Professional Job Roles			
	Data Steward	Data Engineer	Data Scientist	Functional Analyst
Domain Knowledge	3	1	2	3
Data Governance	3	2	2	2
Operational Analytics	3	3	3	3
Data Visualization	2	1	2	3
Research Methods	1	1	3	1
Data Engineering	-	3	1	-
Statistical Techniques	-	1	3	-
Data Methods and Algorithms	-	1	3	-

Computing	1	2	3	1
21st Century Skills	3	3	3	3

Note: Level 1 – basic proficiency; Level 2 – intermediate/average proficiency; Level 3 – advanced/expert-level proficiency

Source: AAP (2018)

7.2. Research Design, Instruments, and Participants

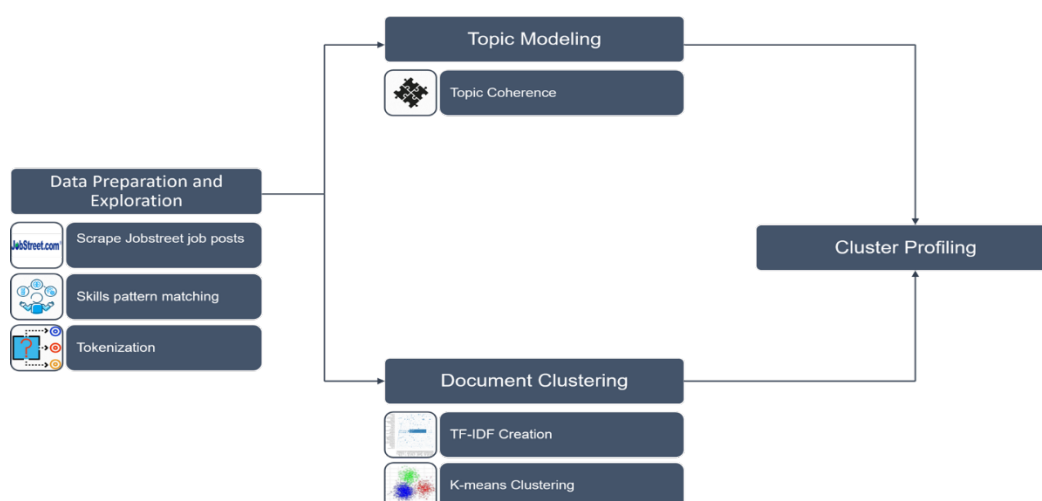
7.2.1. Demand-side

DSA-related activities were identified from the list of activities listed in the Philippine Standard Industry Code (PSIC) of the Philippine Statistics Authority (2009)⁴. The workers who performed these DSA- related activities were assumed to possess the DSA competencies. Employment data from the 2013-2016 Annual Survey of Philippine Business and Industry (ASPBI) on various industries and the 2018 Compilation of Industry Statistics on Labor and Employment were used to determine the number of these workers who possess any of the DSA competencies. Key informant interviews with relevant stakeholders such as those from the Department of Labor and Employment were conducted for additional and supplementary information on the demand for DSA workforce.

The demand for the four (4) DSA job roles and DSA competencies were determined by these tasks: extracting skills and segmenting job roles found in the online job listings (see Appendix A for the detailed online job scraping methodology). The online job scraping methodology followed four stages (Figure 2):

- Stage 1: Data Preparation and Exploration
- Stage 2: Topic Modeling
- Stage 3: Document Clustering
- Stage 4: Cluster Profiling
-

Figure 2. Skills Extraction and Job Segmentation Process



Source: Authors' illustration

⁴ The study manually identified which industries do analytics-related activities are mostly lodged in accordance with the industry definitions provided by PSA, specific in the 2009 Philippine Standard for Industrial Classification (PSIC) and the "analytics" definition used by the framework of this study.

Stage 1: Data Preparation and Exploration.

Scrape Jobstreet job posts. The industry-required skills were sourced from the job descriptions found in these job posts. Job posts from May 1, 2019 to July 21, 2019 were scraped from Jobstreet.com.ph, using keywords that are related to the data profession (e.g. roles, tools, and APEC DSA competencies). The online job scraping lasted for about 3 months and gathered a total of 150, 869 unique job posts. For faster data recovery, 13.6% of these unique job posts (20, 574) were used as a representative sample where the algorithms were performed.

Skills pattern matching and tokenization. Tokenization, the process of segmenting a running text into elements of vocabulary, called words or tokens (Weiss et. al, 2015) was first performed on the extracted skill sets. In an effort to remove the effect of noise, due to inclusion of equal opportunity statements, application instructions, and company benefits, a predetermined set of skills from Wang (2017) was used as vocabulary (see Appendix B). Moreover, to enable classification of these skills into similar groups based on the pattern by which they appear in the job posts, their term frequency-inverse document frequency (TF-IDF) statistics were calculated.

Stage 2: Topic Modeling

Topic modelling, a dimensionality reduction technique which links words of original texts with the same context and discover themes that run through them, called topics, followed afterwards. Non-negative matrix factorization (NMF), a non-probabilistic topic modelling technique was employed in the analysis because it works better and faster when learning keywords or short texts (Cheng et. al, 2013, and Chen et. al, 2016). Since NMF will not automatically provide the optimal number of skills groups, the selection of skills group was based on topic coherence score. This was used in the study since it is the most aligned with human interpretability (Röder et. al, 2015). The extracted 30 skills groups were manually mapped into the 10 APEC-Recommended DSA competencies by select industry experts. Weights per competency of each job post were calculated. This weight signifies the portion of work that requires the competency. The sum of the competency weights should equal to 100%.

Stage 3: Document Clustering

To group the job posts into the four DSA job roles - data steward, data engineer, data scientist and functional analyst-, document clustering was performed. Document clustering is the segmentation of documents in a corpus into homogenous groups with the use of words or terms within each document as features in the model (Manning and Hinrich, 1999). While there are several algorithms to implement clustering, K-means method was the one used in the analysis.

Stage 4: Cluster Profiling

Meetings among field experts were held to refine the groupings from the online job scraping. A survey questionnaire was also used to conduct key informant interviews with industry practitioners and human resource personnel of select IT-BPM companies to gain more insights about the job roles and competencies which employers mostly look for.

7.2.2. Supply-side

The top ten undergraduate degree programs of current DSA practitioners were considered as the ten DSA-related undergraduate degree programs in this study (see Appendix C for the list and count). These DSA practitioners were AAP members or those who attended AAP's monthly meetups.

Academe experts who were either Program Administrators or faculty members of these ten DSA-related undergraduate programs were convened in Focus Group Discussions. The respondents were asked to evaluate two types of curricula of each DSA-related degree program: the actual curriculum of their university⁵ and the sample curriculum found in the respective CHED Memorandum Order (CMO). The actual curriculum represents an innovative program while the sample curriculum represents the minimum requirements for the program.

Each curriculum was evaluated on the extent it is able to equip the graduates with the basic proficiency of the ten analytics competencies following this scale:

- 0 - does not equip** graduates with the analytics competency
- 1 - minimally equips** graduates with the analytics competency
- 2 - moderately equips** graduates with the analytics competency
- 3 - mostly equips** graduates with the analytics competency
- 4 - fully equips** graduates with the analytics competency

Each curriculum was likewise evaluation on its ability to prepare the graduates to perform the four DSA job role – Data Steward, Data Engineer, Data Scientist and Functional Analyst following this scale:

- 0 - does not enable** graduates to perform the tasks of the DSA job role
- 1 - minimally enables** graduates to perform the tasks of the DSA job role
- 2 - moderately enables** graduates to perform the tasks of the DSA job role
- 3 - mostly enables** graduates to perform the tasks of the DSA job role
- 4 - fully enables** graduates to perform the tasks of the DSA job role

Through the evaluations of these experts, the study was able to analyze the supply of DSA workforce.

8. Results and Discussion

8.1. Demand-side

8.1.1. Current availability of DSA competencies

Out of 1,271 activities listed in the Philippine Standard Industry Code (PSIC), twenty-two (22) activities (Table 4) were identified to require any of the 10 DSA competencies. Majority of these DSA related activities belongs to the information and communications subsector (41%), followed by the manufacturing (33%), and professional, scientific and technical activities (24%). The study estimated that, in 2016, there were about 174,000 workers who performed

⁵ The study made use of the curricula from the University of the Philippines Diliman (UPD) and the University of Asia and the Pacific (UA&P). Except for the degree programs Economics, Math/Applied Math and Information Technology which were assessed by UA&P faculty/program directors, all the other identified DSA-related degree programs were assessed by experts of the field in UPD.

these activities and thus, possessed any of the DSA competencies⁶ (Table 4). Such employment size is not far from the estimates provided by APEC Human Resource Technical Working Group on DSA workforce in the Philippines in 2016 which is 147, 420 (see Table 1). This figure was only 0.4 percent⁷ of the total labor force in the Philippines in the said year (see Appendix D for employment share per major industry group). Almost half of these DSA competent workers (i.e., 42%)⁸ belonged to the information technology and business process management (IT-BPM) industry⁹ (Table 4).

Table 4. Employment size of industry subclasses with mostly analytics-related activities

PSIC Industry Code	Industry sub-classes	2016 ¹⁰	
		Employees	%
Manufacturing (33.52%)			
C26200	Manufacture of computers and peripheral equipment and accessories	58,224	33.52%
Financial and Insurance Services (0.11%)			
K66290	Other activities auxiliary to insurance and pension funding	191	0.11%
Professional, Scientific and Technical Activities (24.49%)			
M70200	Management consultancy activities	27,396	15.77%
M71109	Other technical activities related to architectural and engineering	2,902	1.67%
M72101	Research and experimental development in natural sciences	723	0.42%
M72102	Research and experimental development in engineering and technology	342	0.20%
M72103	Research and experimental development in health sciences	465	0.27%
M72104	Research and experimental development in agricultural sciences	257	0.15%

⁶ The researchers made use of the keywords also used in the online job scraping. This is for the purpose of consistency with regard to the definition of Data Science and Analytics used in this study.

⁷ Authors' computations given that there are about 41 million workers in the country in 2016 (Philippine Statistics Authority 2013-2017, Labor Force Survey as cited from the JobsFit 2022 Report of the Bureau of Local Employment, Department of Labor and Employment; PSA CISLE Report, 2018).

⁸ This is also supported by the data gathered through online job scraping. Job postings online can be a reliable source of pinning down current demand for DSA workforce as affirmed in our interviews with some in IT-BPM companies in the country. It is because companies activate all possible channels including online platforms, such as JobStreet, in hiring the right DSA talents, which can reflect actual demand by the IT-BPM industry. Due to lack of available data specific to analytics workforce, the study could not estimate the proportion of these online job postings to the total demand. Moreover, while companies use online platforms to make known their need, they still get to successfully hire needed DSA employees mostly through respective networks and headhunters.

⁹ The Annual Survey of Philippine Business and Industry report identifies business process management BPM industries as determined by two industries, namely, Section J: Information and Communication and Section N: Administrative and Support services (see Appendix E for complete list of industry subclasses under BPM).

¹⁰ See Appendix F for the employment sizes of these industry subclasses for 2013-2015

PSIC Industry Code	Industry sub-classes	2016 ¹⁰	
		Employees	%
M72200	Research and experimental development on social sciences and humanities	1,034	0.60%
M72300	Research and experimental development in information technology	3,686	2.12%
M72400	Research and experimental development services, n.e.c.	s	NA
M73200	Market research and public opinion polling	5,713	3.29%
M749	Other professional, scientific and technical activities, n.e.c.	nad	NA
Information and Communication (41.07%)			
J58200	Software publishing	1,852	1.07%
J62010	Computer programming activities	38,808	22.34%
J62020	Computer consultancy and computer facilities management activities	11,476	6.61%
J62090	Other information technology and computer service activities	3,780	2.18%
J63111	Data processing	14,052	8.09%
J63112	Website hosting services	905	0.52%
J63113	Application hosting services	448	0.26%
Administrative and Support Services (0.83%)			
N82294	Research and analysis activities	291	0.17%
N82299	Other back office operations activities, n.e.c	1,141	0.66%
TOTAL		173,686	100.00%

Source: Philippine Statistics Authority. (2016). *Annual Survey of Philippine Business and Industry (ASPBI)*

8.1.2. Current demand for DSA job roles

Data Science and Analytics (DSA) is at the very nascent stage in the Philippines. The analytics journey begins with convincing leadership that data is the new “oil” of the organization. This challenging task of getting the buy-in of top executives certainly falls on those who operate the business as they can articulate how data can deliver their operational targets. This study revealed that the bulk of the demand for DSA professionals in early 2019 was for the functional analytics role (66% of the job postings) (Table 5).

Once the top management has decided to embrace DSA in the company, the next important DSA job role is the data engineer who ensures that data is available for analytics. The data engineer builds and maintains the right data infrastructures and populates (i.e., extracts, transforms and loads) their central databases with data from transactional systems in the form

ready for model building and analysis. Rightly, the second most sought role in early 2019 is the data engineer (20% of the job postings) (Table 5).

With the clean data in the data warehouses, algorithms can be applied on the data for useful insights. The low demand for data scientist in early 2019 (8% of the job postings) revealed that most companies did not have their data ready for this stage of the data science and analytics process (Table 5).

The Facebook and Cambridge Analytica scandal in 2018 revealed the dark side of analytics. This scandal showed how data of millions of people on Facebook was used without their consent for political advertisement in the 2016 US election and Brexit campaign in UK (MacASkill, 2018). When data is misused, ethical issues arise that violate the human rights of people and put society at risk. Since then, governments have passed or tightened the implementation of data privacy and protection laws. In the Philippines, the Data Privacy Act was passed in 2012 with the National Privacy Commission as the agency mandated to administer and implement it (National Privacy Commission, n.d.). Hence, the demand for data steward has been slightly evident (6% of the job postings) in early 2019 (Table 5). The online scrapping of job postings in early 2019 also revealed that employers sought at least 1-4 years of experience from DSA professionals they want to hire (Table 5).

Table 5. Job profile of in-demand job roles

	DSA Job Roles ¹¹			
	DSt	DE	DS	FA
Collected Number of Job Posts	1, 214	3, 951	1,557	12, 787
% of Job Postings¹²	6.2%	20%	7.9%	65.5%
<i>Less than 1 year experience</i>	5%	10%	19%	9%
<i>1-4 Yrs Experienced Employee</i>	49%	59%	48%	55%
<i>Supervisor/5 Yrs & Up Experienced Employee</i>	28%	24%	20%	17%
<i>Assistant Manager/Manager</i>	15%	4%	12%	14%
<i>Junior Executive</i>	1%	1%	1%	0%
<i>Senior Executive</i>	1%	1%	0%	0%
<i>CEO/SVP/AVP/VP/Director</i>	1%	0%	1%	1%

Source: Authors' Online Job Scraping

8.1.3. Work analysis of in-demand DSA job roles

The work of each of the four job roles was analyzed depending on the extent the competencies, as perceived by employers, are needed to carry out the work (Table 6). Employers believe that the four job roles require all the 10 competencies to be carried out. Research method did not come out as a needed competency because it is usually incorporated with statistical techniques. The 21st century skills are essential for all four job roles. Among the challenges identified by employers when looking for DSA talents is the lack of 21st century skills among graduates.

¹¹ DE – Data Engineers; FA – Functional Analysts; DS – Data Scientists; DSt – Data Stewards

¹² This is the percentage of the total 7, 566 job postings collected online. Only 2, 783 (36.79%) of which are closely related to the three job roles (identified by the Analytics Association of the Philippines Professional Maturity Model)

Companies find it easy to teach programming skills and provide domain experience but not the 21st century workplace skills.

Data governance (27%) and data engineering (28%) are considered by employers to be the two mostly utilized competencies for a data steward. Data engineering, however, is not a competency of a data steward according to the AAP framework. This could be because the task of drafting and implementing data governance policies require thorough understanding of data infrastructure and systems.

Rightly, employers consider data engineering (48%) as the most needed competency to do the job of a data engineer.

The current work of a data scientist equally requires data engineering (16%) and statistical techniques (16%) competencies because he/she must also prepare the data needed for the algorithms to derive useful insights. This is contrary to the AAP framework, which identifies data preparation as a task of the data engineer.

Employers demand that functional analysts possess all the 10 DSA competencies. However, the AAP framework excludes data engineering, statistical techniques and methods and algorithms from the competencies required of functional analysts. This indicates that functional analysts at the workplace also do the technical tasks of the data engineer (i.e. prepare the data) and data scientist (i.e. build models and develop algorithms).

Meanwhile, hiring managers in focus group discussions raised their concern on job candidates not possessing the 21st century skills, the hardest to transmit DSA competency. At the same time, Academics grapple with issues on how to teach this “broad set of knowledge, skills, habits and character traits” in the classrooms (Glossary of Education Forum 2016 par. 1). They admitted that the other nine DSA competencies are much easier to teach.

The discrepancies between employers’ expectations and AAP’s standard definition of the job roles is attributed to the fact that DSA is still at its infancy in most companies in the Philippines where job roles are not delineated and tasks in the data science and analytics process are performed by one person and not a team. This is evident in the responses of hiring managers who said that often they end up hiring the wrong candidate based on incorrect expectations. For instance, while job postings would indicate data science role, the expectation on the job is more of data cleansing and management, which are the primary functions of data engineers. Most members of the recruitment team are not well-versed in the job requirements. As a result, companies experience an average of 46 days delay in hiring the needed DSA job role¹³.

¹³ Information is based on the companies that responded to the survey questionnaire for hiring managers and participants of the Focus Group Discussion among hiring managers on February 2, 2019 at the University of Asia and the Pacific.

Table 6. Work Matrix of DSA Job Roles

Competencies	Data Steward		Data Engineer		Data Scientist		Functional Analyst	
	Employers expectation	Frame work	Employers expectation	Frame work	Employers expectation	Frame work	Employers expectation	Frame work
Domain Knowledge	3%	*	3%	*	4%	*	11%	*
Data Governance	28%	*	8%	*	11%	*	14%	*
Operational Analytics	6%	*	3%	*	10%	*	13%	*
Data Visualization	2%	*	1%	*	11%	*	4%	*
Research Methods		*		*		*		*
Data Engineering	27%	-	48%	*	16%	*	15%	-
Statistical Techniques	1%	-	1%	*	16%	*	3%	-
Methods and Algorithm	3%	-	2%	*	8%	*	4%	-
Computing	12%	*	21%	*	9%	*	6%	*
21st Century Skills	18%	*	10%	*	15%	*	29%	*

Note: This work matrix presents (a) the extent employers expect the competency to be needed to complete the work of the specific job role and (b) whether or not the competency is required by the framework for the specific job role (signifies required).*

Source: Authors' Online Job Scraping & AAP 2018

8. 2. Supply side

8.2.1. DSA-related undergraduate degree programs

A total of thirty (30) Data Science and Analytics-related undergraduate degree programs were identified based on the undergraduate degrees of over a hundred Analytics practitioners who answered the survey questionnaire for this study. The top ten DSA-related degree programs are Computer Science, Business Administration, Statistics, Mathematics or Applied Mathematics, Information Technology, Library and Information Science, Economics, Physics or Applied Physics, Industrial Engineering and Civil Engineering (Table 7).

In 2016, the Bachelor of Science in Business Administration (BSBA) is the most popular DSA-related undergraduate degree program as it is offered by 1,098 (56.5%) HEIs in the country (see Table 7). BSBA is closely followed by BS Information technology, which is offered by 1,089 (56%) HEIs. In the third place is BS Computer Science, which is offered by 756 (38%) HEIs. The accessibility of these programs may explain why most DSA practitioners have any of the top two DSA-related programs as their degree (see Appendix C for specific breakdown).

Table 7. List of Data Science and Analytics-related undergraduate degree programs

Analytics Undergraduate Degree Programs	Number of HEIs offering	% of HEI offering to the total HEIs
1. Computer Science	746 (3 rd)	38.39%
2. Business Administration	1,098 (1 st)	56.51%
3. Statistics	16	0.82%
4. Applied Math/Math	156	8.00%
5. Information Technology	1,089 (2 nd)	56.05%
6. Information Science	74	3.81%
7. Economics	108	5.56%
8. Physics/Applied Physics	25	1.00%
9. Industrial Engineering	109	5.61%
10. Civil Engineering	243	12.51%

Source: Commission on Higher Education (15 March 2019).

The Department of Labor and Employment (DOLE) in its Jobsfit 2022 also identifies the in-demand occupations in 2017 to 2022. All, except Data entry/Encoder Clerk job require at least a college degree. The required undergraduate degrees are mostly these ten DSA-related degrees. All except Data Analyst job require a BS Computer Science (Table 8); giving support to BS Computer Science as the most common degree program of DSA practitioners (see distribution of DSA employees' degree program in Appendix C).

Table 8. In-demand Occupations by Minimum Requirement

Occupation¹⁴ for (2017-2022)	Degree Program
<i>In-demand</i>	
Data Analyst	BS Mathematics BS Statistics
Design Engineering	BS Computer Science
IT Support Staff	BS Information Technology BS Information System BS Computer Science
Software Quality Assurance Analyst	BS Computer Science
Technical Support Specialist	BS Computer Engineering BS Information and Computer Science BS Information Management BS Information Technology
<i>Hard-to-fill</i>	
Computer Hardware Engineer	BS Computer Engineering
Reports Analyst	BS Computer Science BS Statistics BS Business Administration
Search Engine Optimization Analyst	BS Computer Science BS Computer Engineering

¹⁴ Most of these occupations are classified on the basis of the Philippine Standard for Occupational Classification (PSOC) 2012 – with the exception of new or emerging jobs such as Data Scientist, among others.

Occupation ¹⁴ for (2017-2022)	Degree Program
<i>In-demand</i>	
System Analyst	BS Computer Science BS Computer Engineering BS Information Technology
<i>Cross-cutting in-demand and hard-to-fill</i>	
Bioinformatics Specialist	BS Information Technology BS Statistics
Business Analyst	BS Economics BS Business Administration, Banking and Finance
Data Scientist	BS Data Science and Analytics
Information Technology (IT) Specialist	BS Information Technology BS Computer Science BS Information System BS Computer Engineering BS Electrical Engineering
Management Information Analyst	BS Information Technology BS Computer Science BS Information System
Big Data Researcher	BS Statistics BS Business Administration major in marketing BS Data Science and Analytics BS Math BS Economics

Source: Department of Labor and Employment-Bureau of Local Employment (DOLE-BLE). 2017. *JobsFit 2022: Labor Market Information*

8.2.2. Current supply of DSA competencies and DSA-job ready graduates

The curriculum¹⁵ assessments of Program Directors or HEI representatives of the 10 DSA-related undergraduate programs reveal the current supply of DSA competencies and DSA-job ready graduates. Below are the individual program assessments arranged according to their popularity among DSA practitioners. The discussion on the 21st Century Skills is provided in a separate section.

8.2.2.1. Bachelor of Science in Computer Science. The sample CMO curriculum with the minimum requirements of the BS Computer Science prepares its graduates to perform fully the Data Scientist role. The actual curriculum, however, enables its graduates for the Data Engineer role. The evaluation of a few more HEIs supported this assessment that BS Computer Science graduates best perform the tasks of a Data Engineer than a Data Scientist.¹⁶

Since most surveyed DSA practitioners has BS Computer Science degree, it may be expected that this program equips its graduates with all the DSA competencies. However, both actual and sample CMO curricula were assessed to be in need of slight improvement in equipping its

¹⁵ Each recognized undergraduate degree program in the Philippines has an equivalent Commission on Higher Education Memorandum Order (CMO), which is an official document issued by the Commission on Higher Education (CHED). CMOs of degree programs, such as BS Statistics, stipulate policies, standards and guidelines including a sample curriculum for Higher Education Institutions (HEIs) especially the non-autonomous HEIs.

¹⁶ The list of HEIs aside from UPD which were able to assess their respective curriculum on BS CS are: Ateneo de Zamboanga University, Lyceum of the Philippines University, Holy Name University, University of San Carlos, Negros Oriental State University and Manuel S. Enverga University Foundation. Four of these HEIs assessed their BS CS curriculum as able to fully enable graduates to perform the Data Engineer job role.

graduates with the Domain Knowledge competency (Table 9). The sample CMO curriculum further need slight enhancement for Operational Analytics (Table 9).

Table 9. BS Computer Science DSA competencies and job roles

DSA Competencies		Required Competencies based on AAP's framework				BS CS	
		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Domain Knowledge		*	*	*	*	3	3
Data Governance		*	*	*	*	4	4
Operational Analytics		*	*	*	*	4	3
Data Visualization		*	*	*	*	4	4
Research Methods		*	*	*	*	4	4
Data Engineering			*	*		4	4
Statistical Techniques			*	*		4	4
Methods and Algorithm			*	*		4	4
Computing		*	*	*	*	4	4
21 st Century Skills ¹⁷		*	*	*	*		
Degree of enablement to perform >	Actual	0	4	3	0		
	CMO	2	3	4	2		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.2.2. *Bachelor of Science in Business Administration (BS BA)*. Graduates of the actual curriculum of BSBA are ready to perform the Data Steward and Functional Analysts roles (see Table 10). Except for Data Engineering, Statistical Techniques and Data Methods and Algorithms, both DSA job roles require the basic proficiency of other seven (7) DSA competencies (see Table 3). It is then recommended that the sample CMO curriculum on BSBA be significantly improved to enable the students with these seven competencies required for the Data Steward and Functional Analyst roles.

Table 10. BS Business Administration DSA competencies and job roles

Competencies		Required Competencies based on AAP's framework				BS BA	
		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Domain Knowledge		*	*	*	*	4	2
Data Governance		*	*	*	*	1	1
Operational Analytics		*	*	*	*	4	0
Data Visualization		*	*	*	*	3	2
Research Methods		*	*	*	*	3	2

¹⁷ The study discusses only the evaluation of the actual curriculum for the 21st century skills since it was difficult for the participants of the study to accurately assess the CMO sample curriculum of their degree program in terms of the 21st century skills.

Required Competencies based on AAP's framework						BS BA	
Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Data Engineering			*	*		0	0
Statistical Techniques			*	*		4	2
Methods and Algorithm			*	*		4	0
Computing		*	*	*	*	2	1
21 st Century Skills		*	*	*	*		
Degree of enablement to perform >	Actual	4	1	1	4		
	CMO	1	1	1	1		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.2.3. *Bachelor of Science in Statistics*. New Statisticians can immediately work as Data Scientists based on the actual curriculum of BS Statistics. The actual and sample CMO curricula, however, need slight improvement in equipping the students with the basic proficiency of Data Governance as required by the Data Scientist role (see Table 11). Additional minor enhancements must be made to the sample CMO curriculum for Research Methods, Data Engineering, Statistical Techniques and Computing competencies.

Table 11. BS Statistics DSA competencies and job roles

Required Competencies based on AAP's framework						BS STAT	
Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Domain Knowledge		*	*	*	*	4	4
Data Governance		*	*	*	*	3	3
Operational Analytics		*	*	*	*	4	4
Data Visualization		*	*	*	*	4	4
Research Methods		*	*	*	*	4	2
Data Engineering			*	*		4	3
Statistical Techniques			*	*		4	3
Methods and Algorithm			*	*		4	4
Computing		*	*	*	*	4	2
21 st Century Skills		*	*	*	*		
Degree of enablement to perform >	Actual	1	2	4	1		
	CMO	1	1	3	1		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.2.4. *Bachelor of Science in Math/Applied Math*. New Mathematicians are deemed to be enabled to perform the work of a Data Scientist. Since basic proficiency of the ten (10) DSA competencies are required from a Data Scientist (see Table 3), then both actual and sample CMO curricula must be slightly enhanced to better equip the students with competencies in Data Engineering, Methods and Algorithms and Computing (see Table 12).

Table 12. BS Math/Applied Math DSA competencies and job roles

Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	BS Math/Applied Math	
						Actual	CMO
Domain Knowledge		*	*	*	*	4	4
Data Governance		*	*	*	*	4	4
Operational Analytics		*	*	*	*	4	4
Data Visualization		*	*	*	*	4	4
Research Methods		*	*	*	*	4	4
Data Engineering			*	*		3	3
Statistical Techniques			*	*		4	4
Methods and Algorithm			*	*		3	3
Computing		*	*	*	*	3	3
21 st Century Skills		*	*	*	*		
Degree of enablement to perform >	Actual	3	2	4	3		
	CMO	3	2	4	3		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.2.5. *Bachelor of Science in Information Technology (BS IT)*. BSIT graduates are ready to perform as Data Engineers and Functional Analyst (based on the actual curriculum only). However, both curricula need adjustments to equip the students with the basic requirements of the ten competencies as required for the Data Engineer role (see Table 3). The actual curriculum needs slight improvements for Operational Analytics and major improvements for Statistical Techniques and Methods and Algorithms (see Table 13). The sample CMO curriculum, on the other hand, needs slight improvements for Data Governance, Research Methods, and Computing and major enhancements for Domain Knowledge, Operational Analytics, Methods and Algorithms. Moreover, the sample CMO curriculum must incorporate Data Visualization (see Table 13).

Looking into other responses of participating universities, it is clear that between a Data Engineer and Functional Analyst, BS IT graduates are mostly enabled to perform the Functional Analyst role¹⁸. Nonetheless, the “mostly enable” evaluation of the degree program on other job roles also point to the capacity of the program to hone graduates in other job roles.

¹⁸ The list of HEIs aside from UA&P which were able to assess their respective curriculum on BS CS are: Ateneo de Zamboanga University, St. Mary's University, Holy Name University, Negros Oriental State University and Nueva Ecija University of Science and Technology. Four of these HEIs assessed their BS CS curriculum as able to fully enable graduates to perform the Data Engineer job role.

Table 13. BS Information Technology DSA competencies and job roles

Required Competencies based on AAP's framework						BS IT	
Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Domain Knowledge		*	*	*	*	4	1
Data Governance		*	*	*	*	4	3
Operational Analytics		*	*	*	*	3	1
Data Visualization		*	*	*	*	2	0
Research Methods		*	*	*	*	4	3
Data Engineering			*	*		4	4
Statistical Techniques			*	*		2	4
Methods and Algorithm			*	*		2	2
Computing		*	*	*	*	4	3
21 st Century Skills		*	*	*	*		
Degree of enablement to perform >	Actual	3	4	2	4		
	CMO	2	4	1	2		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.2.6. *Bachelor of Library and Information Science (BLIS)*. Based on the actual curriculum, graduates of BSIS are ready to perform the Data Engineer role. However, this curriculum needs slight enhancements to fully equip the graduates with all the ten DSA competencies (see Table 14). The sample CMO curriculum needs major enhancements to provide the students with the basic proficiency of all the DSA competencies as required by the Data Engineer role (see Table 3).

Table 14. BS Library and Information Science DSA competencies and job roles

Required Competencies based on AAP's Framework						BLIS	
Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Domain Knowledge		*	*	*	*	2	2
Data Governance		*	*	*	*	3	2
Operational Analytics		*	*	*	*	2	1
Data Visualization		*	*	*	*	3	1
Research Methods		*	*	*	*	3	2
Data Engineering			*	*		3	3
Statistical Techniques			*	*		3	2
Methods and Algorithm			*	*		3	2
Computing		*	*	*	*	3	1

Required Competencies based on AAP's Framework						BLIS	
Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
21st Century Skills		*	*	*	*		
Degree of enablement to perform >	Actual	2	4	1	3		
	CMO	2	2	1	2		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.2.7. Bachelor of Science in Physics/Applied Physics. The actual curricula of BS Physics and Applied Physics mirrors their counterpart sample CMO curriculum since the sample CMO curriculum for BS Physics/Applied Physics was taken from the actual curricula used in this study¹⁹. New Physicists are deemed capable of fully assuming the Data Scientist role (see Table 15). However, both curricula need slight enhancements for Data Visualization and Methods and Algorithms competencies and major enhancements for Data Governance, Operational Analytics, and Computing competencies (see Table 15) to better equip their students with the basic proficiencies of the DSA competencies as required from a Data Scientist (see Table 3).

Table 15. BS Physics/Applied Physics DSA competencies and job roles

Required Competencies based on AAP's Framework						BS Physics/Applied Physics	
Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Domain Knowledge		*	*	*	*	4	4
Data Governance		*	*	*	*	2	2
Operational Analytics		*	*	*	*	2	2
Data Visualization		*	*	*	*	3	3
Research Methods		*	*	*	*	4	4
Data Engineering			*	*		4	4
Statistical Techniques			*	*		4	4
Methods and Algorithm			*	*		3	3
Computing		*	*	*	*	2	2
21st Century Skills		*	*	*	*		
Degree of enablement to perform >	Actual	2	2	4	3		
	CMO	2	2	4	3		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.2.8. Bachelor of Science/Bachelor of Arts in Economics. Minimal difference can be noted between the assessment of the degree program's actual and sample CMO curricula (see Table 16). The job role that this degree program fully enables its graduates to perform is

¹⁹ According to the Respondent for the BS Physics/Applied Physics degree program

the Functional Analyst. For this role, both curricula need major improvements to fully equip the graduates with the basic proficiency of the competencies required from a Functional Analytics (see Table 3). These competencies are Domain Knowledge, Data Governance, Operational Analytics, Data Visualization, Research Methods, and Computing.

Table 16. BS Economics DSA competencies and job roles

Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	BS Econ	
						Actual	CMO
Domain Knowledge		*	*	*	*	2	2
Data Governance		*	*	*	*	0	0
Operational Analytics		*	*	*	*	2	2
Data Visualization		*	*	*	*	2	1
Research Methods		*	*	*	*	3	3
Data Engineering			*	*		0	0
Statistical Techniques			*	*		3	3
Methods and Algorithm			*	*		3	3
Computing		*	*	*	*	0	0
21 st Century Skills		*	*	*	*		
Degree of enablement to perform >	Actual	1	1	3	4		
	CMO	1	1	3	4		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.2.9. *Bachelor of Science in Industrial Engineering (BS IE)*. The graduates of BSIE are able to immediately assume the DSA job roles of Data Steward, Data Engineer and Functional Analyst. They require additional training to perform the tasks of a Data Scientist. Its actual curriculum only needs slight improvements on Data Governance, and Methods and Algorithms. The sample CMO curriculum, on the other hand, requires slight improvements on all the competencies except for Domain Knowledge (see Table 17).

Table 17. BS Industrial Engineering DSA competencies and job roles

Competencies		Required Competencies based on AAP's framework				BS IE	
		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Domain Knowledge		*	*	*	*	4	4
Data Governance		*	*	*	*	3	3
Operational Analytics		*	*	*	*	4	3
Data Visualization		*	*	*	*	4	3
Research Methods		*	*	*	*	4	3
Data Engineering			*	*		4	3
Statistical Techniques			*	*		4	3
Methods and Algorithm			*	*		3	3

Required Competencies based on AAP's framework						BS IE	
Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Computing		*	*	*	*	4	3
21 st Century Skills		*	*	*	*		
Degree of enablement to perform >	Actual	4	4	3	4		
	CMO	4	4	3	4		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.2.10. *Bachelor of Science in Civil Engineering (BS CE)*. Among the ten DSA-related programs, BSCE program is least fitted for the DSA job requirements. New Civil Engineering graduates are not fully prepared to take on any DSA job role. At most, they can be hired as Data Scientists subject to additional training.

Graduates of BSCE do not possess three DSA competencies, namely Data Governance, Operational Analytics, and Data Engineering. Based on the actual curriculum, graduates only have three DSA competencies, namely Research Methods, Statistical Techniques and Computing. Moreover, the sample CMO curriculum barely equip them with the other five DSA competencies (see Table 18).

Table 18. BS Civil Engineering DSA competencies and job roles

						BS CE	
Competencies		Data Steward	Data Engineer	Data Scientist	Functional Analyst	Actual	CMO
Domain Knowledge		*	*	*	*	2	1
Data Governance		*	*	*	*	0	0
Operational Analytics		*	*	*	*	0	0
Data Visualization		*	*	*	*	2	2
Research Methods		*	*	*	*	4	1
Data Engineering			*	*		0	0
Statistical Techniques			*	*		4	1
Methods and Algorithm			*	*		2	0
Computing		*	*	*	*	4	2
21 st Century Skills		*	*	*	*		
Degree of enablement to perform >	Actual	0	0	3	2		
	CMO	0	0	3	0		

Source: Survey responses of participating Program Directors or HEI representatives

8.2.3. DSA-related Degree Programs and the 21st Century Skills

A key feature among the APEC-Recommended Data Science and Analytics competencies is the 21st Century Skills. This competency is often contrasted with the other competencies as it pertains to as set of non-technical skills. Hiring managers indicated that such skills are the most difficult to develop and most valuable in the workplace. Hence, the discussion on the 21st Century Skills merits a special section.

Employers expect that students acquire these 21st Century Skills while gaining their undergraduate degrees²⁰. It is then important to determine the extent HEIs are able to incorporate these 21st Century Skills in the curricula of the DSA-related degree programs. As previously mentioned, 21st Century Skills refer to a set of sixteen (16) “broad set of knowledge, skills, habits and character traits” (Glossary of Education Forum 2016 par. 1) otherwise called “soft skills”.

Assessment of the actual curricula of the ten DSA-related programs reveal that their graduates possess these sixteen soft skills in varying degrees: critical thinking and problem solving skills (9 of 10 DSA-related programs); Job Collaboration, Communication and Storytelling and Decision-making (8 of 10 DSA-related programs); Planning and Organizing skills and Job Flexibility (7 of 10 DSA-related programs); Ethical Mindset and Organizational Awareness (5 of 10 DSA-related programs); Customer Focus, Social Awareness and Professional Networking (4 of 10 DSA-related programs); Business Fundamentals and dynamic (self) re-skilling, (3 of 10 DSA-related programs); Cross-cultural Awareness (2 of 10 DSA-related programs) and Entrepreneurship (1 of 10 DSA-related programs) (Table 19).

Table 19. 21st century skills component of DSA-related undergraduate programs’ actual curricula

21st Century Skills	DSA-related undergraduate program									
	BS CS	BS BA	BS Stat	BS Math/ Applied Math	BS IT	BLIS	BS Econ	BS Physics/ Applied Physics	BS IE	BS CE
Job Collaboration	4	4	4	4	4	3	4	4	4	2
Communication and Storytelling	4	4	4	4	4	3	4	4	4	3
Ethical Mindset	4	3	4	4	4	3	4	2	3	2
Organizational Awareness	4	4	3	4	4	3	2	1	4	2
Critical Thinking	4	4	4	4	4	4	4	4	4	3
Planning and Organizing	4	4	4	4	4	4	3	3	4	2
Problem Solving	4	4	4	4	4	4	3	4	4	4
Decision Making	4	4	4	4	4	4	3	4	4	3
Customer Focus	3	4	2	3	4	4	2	1	4	2
Job Flexibility	4	4	3	4	4	4	3	4	4	2
Business Fundamentals	3	4	2	2	4	2	2	0	4	2
Cross-cultural Awareness	4	3	2	3	4	3	2	1	3	1
Social Awareness	4	4	3	3	4	4	3	1	3	1
Dynamic (self) re-skilling	4	3	3	2	4	4	2	2	3	3

²⁰ Insights from the Focus Group Discussion with hiring managers for DSA job roles on February 2, 2019 at the University of Asia and the Pacific

21st Century Skills	DSA-related undergraduate program									
	BS CS	BS BA	BS Stat	BS Math/ Applied Math	BS IT	BLIS	BS Econ	BS Physics/ Applied Physics	BS IE	BS CE
Professional Networking	3	4	2	3	4	4	4	3	3	2
Entrepreneurship	3	3	1	3	4	2	1	0	3	3

Source: Survey responses of participating Program Directors or HEI representatives

Assessing the sample CMO curricula of the nine DSA-related programs²¹ reveal that their graduates possess these eighteen soft skills in varying degrees: problem solving skills (7 of 9 DSA-related programs); Job Collaboration and Critical Thinking (5 of 9 DSA-related programs); Organizational Awareness, Planning and Organizing skills and Decision-making (4 of 9 DSA-related programs); Communication and Storytelling and Job Flexibility (3 of 9 DSA-related programs); Ethical Mindset and Customer Focus (3 of 9 DSA-related programs); Business Fundamentals, Cross-cultural Awareness and Dynamic (self) Re-skilling (1 of 9 DSA-related programs); Social Awareness, Professional Networking and Entrepreneurship (0 of 9 DSA-related programs) (Table 20).

Overall, these DSA-related programs can be improved specifically on seven “soft skills”: Entrepreneurship, Professional Networking, Social Awareness, Dynamic (self) Re-skilling, Cross-cultural Awareness, Business Fundamentals and Customer Focus.

Among the sample CMO curricula, the BS Industrial Engineering program best educates its graduates with 11 of 16 “soft skills”; BS Math and Applied Math with 9 of 16 “soft skills”; BS Computer Science with 8 of 16 “soft skills”; BS Physics and Applied Physics with 6 of 16 “soft skills”; BLIS with 5 of 16 “soft skills”; BS Statistics with 2 of 16 “soft skills”; BS Civil Engineering with 1 of 16 “soft skills”; BS Business Administration and BS Information Technology with 0 of 16 “soft skills” (see Table 20).

Overall, BS Business Administration, BS Information Technology and BS Statistics can be improved significantly by incorporating most of the 16 “soft skills” in their respective sample CMO curricula.

Table 20. 21st century skills component of DSA-related undergraduate programs’ CMO curricula

21st Century Skills	DSA-related undergraduate program (based on CMO's sample curriculum) ²²								
	BS CS	BS BA	BS Stat	BS Math/ Applied Math	BS IT	BLIS	BS Physics/ Applied Physics	BS IE	BS CE
	4	2	4	4	3	2	4	4	2
	3	2	3	4	1	2	4	4	3
	4	3	3	4	2	2	2	3	3

²¹ Respondent opted not to assess the sample CMO curriculum for BS Economics.

²² Identified respondent for evaluating sample curriculum of BS Economics opted to not answer due to objective reasons.

Organizational Awareness	4	3	2	4	1	4	1	4	2
Critical Thinking	4	2	3	4	3	4	4	4	2
Planning and Organizing	4	3	3	4	2	4	3	4	2
Problem Solving	4	2	4	4	3	4	4	4	4
Decision Making	4	2	3	4	2	3	4	4	3
Customer Focus	2	3	1	3	1	4	1	4	1
Job Flexibility	2	3	2	4	1	3	4	4	2
Business Fundamentals	1	3	1	2	1	2	0	4	1
Cross-cultural Awareness	2	2	1	3	0	3	1	4	2
Social Awareness	2	3	1	3	3	3	1	3	2
Dynamic (self) re-skilling	4	2	1	2	0	3	2	3	2
Professional Networking	2	3	1	3	0	3	3	3	2
Entrepreneurship	2	3	1	3	0	2	0	3	2

Source: Survey responses of participating Program Directors or HEI representatives

Additional insights can be drawn from the findings of the 2018 Philippine Talent Map Initiative (PTMI)²³ (see Table 21) which profiled the Philippine labor force on a set of “soft skills”. The “soft skills” that are included in both PTMI and AAP’s 21st Century skills that were assessed to be poor among Filipino workers are decision-making, planning and organizing and problem-solving skills. This findings of PTMI are consistent with the results of this study insofar students of DSA-related degree programs are equipped with these 21st Century skills. As previously discussed, innovative DSA-related programs adequately incorporate these three “soft skills” in their actual curricula. However, those that simply adopt the minimum requires of CHED often miss-out these “soft skills” in their curricula²⁴. Since most HEIs in the Philippines follow the sample CMO curriculum, it can be assumed that many students and job starters have poor training on these 21st Century skills. So, this study recommends that the sample CMO curriculum for these DSA-related degree programs be improved to incorporate these 21st Century skills.

²³ A Department of Labor and Employment (DOLE) and SFI Group Joint Initiative: Profiling the 21st Century Skills of the Philippine Labor Force in 2018

²⁴ Number of DSA-related curricula that incorporates the “soft skills”: Decision-making (8/10 of actual curriculum and 4/9 of sample CMO curriculum); Planning and Organizing (7/10 of actual curriculum and 4/9 of sample CMO curriculum); Problem-solving (9/10 of actual curriculum and 7/9 of sample CMO curriculum)

Table 21. Mean Scores of the Philippines for 21st Century Skills

■ PHILIPPINES OVERALL MEAN SCORES		■ STUDENTS OVERALL MEAN SCORES	
ENGLISH FUNCTIONAL SKILL	72	ENGLISH FUNCTIONAL SKILL	74
MATH FUNCTIONAL SKILL	72	MATH FUNCTIONAL SKILL	73
WORKPLACE ETHICS	72	WORKPLACE ETHICS	72
MULTI-TASKING	70	ENGLISH COMPREHENSION	71
ENGLISH COMPREHENSION	69	MULTI-TASKING	70
CRITICAL THINKING	66	CRITICAL THINKING	68
STRESS TOLERANCE	65	STRESS TOLERANCE	65
SOCIAL PERCEPTIVENESS	64	SOCIAL PERCEPTIVENESS	65
SELF MOTIVATION	61	PROBLEM SENSITIVITY	62
TEAMWORK	61	SELF MOTIVATION	62
PROBLEM SENSITIVITY	61	TEAMWORK	62
DECISION MAKING	57	INNOVATION	59
INNOVATION	57	DECISION MAKING	59
PLANNING AND ORGANIZING	57	PLANNING AND ORGANIZING	57
CREATIVE PROBLEM SOLVING	54	CREATIVE PROBLEM SOLVING	56

■ TRAINEE OVERALL MEAN SCORES	
ENGLISH FUNCTIONAL SKILL	62
WORKPLACE ETHICS	61
SOCIAL PERCEPTIVENESS	60
MULTI-TASKING	60
STRESS TOLERANCE	59
MATH FUNCTIONAL SKILL	58
ENGLISH COMPREHENSION	57
CRITICAL THINKING	55
SELF MOTIVATION	52
TEAMWORK	52
PROBLEM SENSITIVITY	50
CREATIVE PROBLEM SOLVING	49
DECISION MAKING	48
PLANNING AND ORGANIZING	48
INNOVATION	45

Source: Department of Labor and Employment and SFI Group of Companies (DOLE and SFI Group). (2018). *Philippine Talent Map Initiative: Profiling the 21st Century Skills of the Philippine Labor Force*.

8.2.4. Summary of Supply-side discussion

The curriculum²⁵ assessments of HEI representatives revealed that the 10 DSA- related undergraduate programs equip their graduates with the basic proficiency of the 10 DSA competencies in varying degrees (Table 22). Among these programs, the BS Computer Science program performs best in fully enabling its students with 7 of 10 DSA competencies. Graduates in Mathematics may be expected to possess 6 of 10 DSA competencies. BS IT and BS Economics graduates possess only 2 of the DSA competencies; Industrial Engineering only 1 of 10 DSA competencies. Surprisingly, programs in business, information science and civil engineering are inadequate in educating their students with any of the DSA competencies. This is an indication that current DSA practitioners with these undergraduate degrees gained these DSA competencies through other means.

Among the 10 DSA competencies, Statistical Techniques is effectively taught by half of the DSA-related programs. Domain Knowledge and Data Engineering are taught effectively in 4 of the 10 DSA-related programs. The 21st century skills are the hardest to transmit to students.

²⁵ Each recognized undergraduate degree program in the Philippines has an equivalent Commission on Higher Education Memorandum Order (CMO), which is an official document issued by the Commission on Higher Education (CHED). CMOs of degree programs, such as BS Statistics, stipulate policies, standards and guidelines including a sample curriculum for Higher Education Institutions (HEIs) especially the non-autonomous HEIs.

In 2016, the most available competency among the graduates is Statistical Techniques, followed by Data Engineering. It would be hard to find someone with 21st century skills, Operational Analytics or Domain Knowledge.

Table 22: Graduates equipped with DSA competencies

DSA-related program	BS CS	BS BA	BS STAT	BS MATH	BS IT	BLIS	BS ECON	BS PHYSICS	BS IE	BS CE	Total
Competencies											
Domain Knowledge			4	4				4	4		6,026
Data Governance	4			4							13,331
Operational Analytics			4	4							2,573
Data Visualization	4		4	4							13,688
Research Methods	4			4				4			13,577
Data Engineering	4				4		4	4			54,681
Statistical Techniques	4			4	4		4	4			56,897
Methods and Algorithms	4		4								11,472
Computing	4										11,115
21 st Century Skills											
											157, 149

Note: a.) "4" indicates that the undergraduate program fully enables its students with the specified DSA competency; b.) Figures for the total number of graduates per DSA-related degree program (SY: 2016-2017) were taken from CHED (2019): BS CS = 11,115; BSBA = 84,894; BS STAT = 357; BS Math = 2216; BS IT = 42,039; BS Econ = 548; BS Physics = 1,281; BS IE = 246; BS CE = 3,207

Source: Survey responses of participating Program Directors or HEI representatives; Authors' computations; CHED 2019

The curriculum assessment further showed that in 2016 there were 68,518 DSA job ready graduates from the 10 DSA related undergraduate programs (Table 23). A total of 45,246 (68%) of them were ready to be data engineers with degrees in Information Technology and Industrial Engineering. Meanwhile, 13,577 (or 20%) of these graduates were ready to work as data scientists with degrees in Computer Science, Mathematics and Physics. Significantly fewer number of graduates were fully enabled to perform the functional analyst (4, 488 or 7%) and data steward (3,207 or 5%) roles (Table 5).

Table 23. Job-ready graduates from DSA-related undergraduate programs

DSA-related degree program	BS CS	BS BA	BS STAT	BS MATH	BS IT	BLIS	BS ECON	BS PHYSICS	BS IE	BS CE	Total
Job roles											
Data Steward									4	x	3,207
Data Engineer					4				4	x	45,246
Data Scientist	4			4				4			13,577
Functional Analyst							4		4	x	4,488
											66,518

Note: a.) "X" indicates that the undergraduate program does not prepare its students to perform the indicated job role.

Source: Survey responses of participating Program Directors or HEI representatives; Authors' computations; CHED 2019

Comparing the distribution of demand and supply of the DSA job roles (Table 24), the study found evidence of misalignment between the type of DSA workers sought by employers and the type of DSA graduates educated by our higher education institutions. There is an oversupply of data engineers and data scientists and an undersupply of functional analysts and data steward. Employers are looking mostly for functional analysts (66% of the DSA demand) while higher education institutions are mostly graduating job-ready data engineers (68% of the DSA supply).

Table 24. Distribution for Demand and Supply of DSA Job Roles

	Data Steward	Data Engineer	Data Scientist	Functional Analyst
Demand	8%	20%	6%	66%
Supply	5%	68%	20%	7%

Note: Values may not add up to totals due to rounding of figure

Sources: (a) For Demand: online job scrapping and (b) for Supply: survey responses of participating Program Directors or HEI representatives.

9. Conclusion: Demand and Supply

The findings indicate that there is a misalignment between the demand and supply of DSA workforce in the country. This gap is revealed based on the job roles demanded by employers and supplied by educators in 2016. While employers look for graduates who are fully-enabled to perform the work of a functional analyst, higher education produces graduates who are apt to be data engineers. Consistently, in terms of competencies, DSA-related degree programs are able to fully equip its graduates with competencies mostly required to perform tasks of a data scientist or data engineers (statistical techniques and data engineering) than those competencies needed for the data steward or functional analyst job roles (domain knowledge and operational analytics).

Such misalignment is mostly attributed to the fact that DSA is still at its infancy in most companies in the Philippines. Hence, job roles are often not yet specifically distinguished based on the DSA process and so performed only by one person rather than a team. Employers then

also run the risk of not looking for the right worker for the right set of tasks. This is evident in the discrepancies between employers' expectations from their newly-hires and AAP's definition of the DSA job roles. At the same time, curricula of DSA-related degree programs can still be improved in line with the emerging demand for DSA competencies.

With this misalignment between demand and supply of DSA competencies comes the impetus for the government to work on appropriate mechanisms. If not, youth unemployment may exacerbate.

10. Recommendations

The findings of the study provided baseline evidences on the current DSA skills shortage. They disclosed that the supply of DSA competencies in the country is not enough to meet the current demand of the industry although existing degree programs fully equip graduates with the DSA competencies required by employers.

Given the results and limitations of the study, two main recommendations can be considered towards building an adequate supply of DSA workers in the country: a.) the use of AAP's framework to define the Data Science and Analytics profession, and b.) the promotion of industry-academe-government linkages that can help analytics activities in the country mature and expand into an industry and contribute more to the economy.

10.1. Use of AAP's Framework

Various strategies have been presented in relevant literature and recent conferences about the Fourth Industrial Revolution (FIRe). However, many of these strategies can successfully take off by first having a common understanding of this emerging market for DSA workforce.

Other economies have already begun such initiative and have been making huge steps in developing their respective DSA workforce. Examples of these initiatives are EDISON Data Science framework in Europe, BHEF in the United States and SkillsFuture in Singapore. Indeed, a framework that can facilitate a common understanding among various stakeholders can help in aligning the demand and supply of DSA workforce in the country. Currently, AAP's professional maturity model is continuously being enhanced and so can be an appropriate starting point especially in profiling the demand and supply of DSA competencies and job roles in the country.

1. In terms of demand, analytics companies or associations can make use of the framework as basis for hiring and developing DSA talents. Moreover, relevant government agencies such as the Department of Trade and Industry (DTI) and the Department of Information, Communication and Technology (DICT) can also collaborate more efficiently with industry players through policies and programs fitted to the industry needs.
2. In terms of the supply, CHED can make use of the framework in creating new standards for Data Science and Analytics degrees or updating existing CMOs of DSA-related degree programs to accommodate the DSA industry requirements.

HEIs can likewise use the framework to improve their existing DSA-related degree programs to increase their capacity in enabling the graduates to be industry-ready

DSA workers. In a specific way, they can incorporate courses that particularly develop DSA competencies and aim to enable their graduates to perform any of the DSA job roles. In so doing, these degree programs can better educate more enabled DSA workforce. For instance, BS Statistics programs educate Statisticians who are industry-ready to perform the tasks of a Data Scientist. This curriculum updating can be a potential solution to augment the current DSA skills shortage in the short run.

HEIs can likewise use the framework to develop and offer undergraduate programs specially aimed to graduate Data Stewards, Data Engineers, Data Scientists and Functional Analysts with DSA competencies.

3. This study puts forward the use of a common understanding of analytics. However, it covered only the demand of the IT-BPM sector and the supply of DSA workforce from 10 DSA-related undergraduate programs. Hence, this study recommends that this scope be expanded to capture a more inclusive picture of the demand and supply of DSA workforce in the country. Future studies can include more sectors and more undergraduate programs. In addition, they also consider other sources of supply of DSA workforce such as the online platforms (ex. Coursera, Moot, etc.) or training centers that offer DSA courses.

10.2. Other industry-academe-government linkages

A common definition of Data Science and Analytics profession can facilitate industry-academe-government linkages aimed to align the demand and supply of DSA workforce.

1. Academe and government, specifically through CHED, can work together in establishing standards for degree programs in data science and analytics or updating existing CMOs of degree programs.
2. The Academe can involve industry in improving curriculum, course design and delivery of instructions or in co-developing teaching materials.
3. Industry players can encourage their own employees to teach and impart a practical background of DSA to students.
4. Companies can share the cost of DSA education, such as data laboratories, hardware, and software licenses.

Industry players can help intensify faculty development through faculty exchange programs and research collaborations aimed to increase teachers' theoretical and practical background of the field. An example of this enrichment that is currently being done is the Data Incubator Graduate Program of the Binghamton University. The program is an intensive 8-week fellowship that prepares master's students, PhDs, and postdocs in STEM and social science fields for industry careers as Data Scientists. But instead of fishing them out of the universities, companies can partner with the universities for the expertise of their professors.

5. The government working closely with the industry and academe can play a major role in formulating policies and programs that promote DSA. The government's most relevant agencies are Department of Trade and Industries (DTI) to facilitate the evolution of analytics as an industry, Department of Information and Communication Technology (DICT) to enable the needed information and communication technology infrastructure for DSA, Department of Labor and Employment (DOLE) to facilitate the healthy balance between the demand and supply of DSA workforce, Department of Science and Technology (DOST) to encourage the proliferation of research projects in DSA, Philippine Statistical Authority (PSA) to include DSA in its databases and studies and Commission on Higher Education (CHED) to promote DSA education among the Higher Education Institutions (HEIs).

REFERENCES

Aggarwal, C. and Zhai, C. (2012). *Mining Text Data*. New York, NY: Springer.

Analytics Association of the Philippines (AAP). 2018. 10 analytics competencies. Taguig City, Philippines: AAP.

APEC Human Resource Development Working Group. 2017. Data science and analytics skills shortage: Equipping the APEC workforce with the competencies demanded by employers. Singapore, Singapore: APEC Secretariat.
<https://www.apec.org/Publications/2017/11/Data-Science-and-Analytics-Skills-Shortage> (accessed on August 11, 2018)

APEC Human Resources Development Working Group. 2019 July 19. Close the Digital Skills Gap by 2025 through Collaboration. Singapore. Retrieved from
https://www.apec.org/Press/News-Releases/2019/0719_Digital (accessed August 1, 2019)

Arthur, D. and Vassilvitskii, S. (2007). *k-means++: The advantages of careful seeding*. Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics.

Business Higher-Education Forum (BHEF). (2014). *Data Science Emerges as Essential Tool for Decision Making and Innovation Across Industry Sectors*. Business Higher-Education Forum.

Commission on Higher Education. Graduates by Specific Baccalaureate Programs_2013-2016 as of 15 March 2019. *Unpublished (data request)*

Chen, Y., Bordes, J., and Filliat, D. (2016). *An experimental comparison between NMF and LDA for active cross-situational object-word learning*. Sixth Joint IEEE International Conference Developmental Learning and Epigenetic Robotics (ICDL-EPIROB), Cergy-Pontoise, France, Sep 2016.

Cheng, X., Guo, J., Liu, S., Wang, Y., & Yan, X. (2013). *Learning Topics in Short Texts by Non-negative Matrix Factorization on Term Correlation Matrix*. SDM.

Elemia, C. (2016, September 17). Villanueva: We lack graduates who match jobs. Retrieved from Rappler: <https://www.rappler.com/nation/146475-villanueva-we-lack-graduates-match-jobs> (accessed August 12, 2018)

Dadios et al.(2018). Preparing the Philippines for the Fourth Industrial Revolution: A Scoping Study: A Scoping Study. *Discussion Paper Series No. 2018-11*. Retrieved from <https://pidswebs.pids.gov.ph/CDN/PUBLICATIONS/pidsdps1811.pdf>

Department of Labor and Employment-Bureau of Local Employment (DOLE-BLE). 2017. *JobsFit 2022: Labor Market Information*. Retrieved from <http://www.ble.dole.gov.ph/downloads/Jobsfitepublications/DOLE-BLE-JobsFit2022.pdf>

Department of Labor and Employment and SFI Group of Companies (DOLE and SFI Group). (2018). *Philippine Talent Map Initiative: Profiling the 21st Century Skills of the Philippine Labor Force*.

Demchenko, Y. (2017). *EDISON Data Science Framework: Part 4. Data Science Professional Profiles (DSPP) Release 2*. Education for Data Intensive Science to Open New science frontiers (EDISON) Project. Working Document. Retrieved from https://www.kfa-juelich.de/SharedDocs/Downloads/UE/DE/Controlling/Workshop_WiDW/17-07-07_Demchenko.pdf?__blob=publicationFile

Francisco et. al. (2019). Mapping Philippine Workers at Risk of Automation in the Fourth Industrial Revolution. *Working Paper 2019-01*. Asian Institute for Management Rizalino S. Navarro Policy Center for Competitiveness. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3366809

MacASkill, A. 2018. What are the links between Cambridge Analytica and a Brexit campaign group?. <https://www.reuters.com/article/us-facebook-cambridge-analytica-leave-eu/what-are-the-links-between-cambridge-analytica-and-a-brexit-campaign-group-idUSKBN1GX2IO> (accessed on December 16, 2019)

National Privacy Commission, n.d. A brief primer on republic act 10173. <https://www.privacy.gov.ph/data-privacy-act-primer/> (accessed on December 16, 2019)

Manning, C. and Hinrich, S. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

McKinsey Global Institute. (2016). *People on the Move: Global Migration's Impact and Opportunity*. McKinsey&Company.

Microsoft. (2017 February 6). *Unlocking the Economic Impact of Digital Transformation in Asia Pacific: Microsoft Asia Digital Transformation Study 2018 - Southeast Asia and the Philippines*.

Philippine Statistics Authority. (2009). Philippine Standard for Industrial Classification (PSIC). Retrieved from https://psa.gov.ph/sites/default/files/PSA_PSIC_2009.pdf

Philippine Statistics Authority. (2012). Philippine Standard for Occupational Classification (PSOC). Retrieved from <https://psa.gov.ph/sites/default/files/4publication.pdf>

Philippine Statistics Authority. (2016). *Annual Survey of Philippine Business and Industry (ASPBI)*. Retrieved from <https://psa.gov.ph/statistics/survey/business-and-industry/index>

Philippine Statistics Authority. (2018). *Compilation of Industry Statistics on Labor and Employment*. Retrieved from <https://psa.gov.ph/sites/default/files/2018%20CISLE%20Final.pdf>

Pricewaterhouse Cooper and Business-Higher Education Forum (PwC and BHEF). (2017). *Investing in America's data science and analytics talent: the case for action*. Retrieved from http://www.bhef.com/sites/default/files/bhef_2017_investing_in_dsa.pdf

Ramos, C. (2016 May). Skills Issues, Sources of Skills Issues, and Policy Responses in Five ASEAN- Member Countries: Indonesia, the Philippines, Thailand, and Vietnam. The HEAD Foundation

Röder, M., Both, A., and Hinneburg, A. (2015). *Exploring the Space of Topic Coherence Measures*. Shanghai, China: ACM.

Schutt, R. & O’Neil, C. 2013. Doing data science: straight talk from the frontline. O’Reilly Media Inc. United States of America

Tableau. (2016). *The State of Data Education in 2016: How U.S. higher education responds to the data skills gap*. Retrieved from <https://www.tableau.com/data-education-2016>

Tan, K.S. & Tang, J. (2016). *New skills at work: Managing skills challenges in ASEAN-5*. Research Collection School of Economics. Singapore Management University.

Tan & Tan. n.d. Need For B.S. Data Science Degree Program in the Philippines. Philippine Statistics Authority. Retrieved from <https://psa.gov.ph/sites/default/files/Session%203-8%20Need%20for%20B.S.%20Data%20Science%20Degree%20Program.pdf>

Wang, W. (2017, August 19). *Job Skills Extraction*. Retrieved from <https://github.com/2dubs/Job-Skills-Extraction>.

Weiss, S., Indurkha, N., and Zang, T. (2015). *Fundamentals of Predictive Text Mining*. London: Springer-Verlag.

Xu, R. (2005). Survey of Clustering Algorithms. *IEEE Transactions on Neural Networks*, Vol. 16, No. 3, May 2005.

APPENDICES

Appendix A. Detailed Online Job Scraping Methodology

- **Definition 1: Tokenization** is the process of segmenting a running text into elements of vocabulary, called words or tokens. [Weiss et. all (2015)]
- **Definition 2: Topic Modeling** is a dimensionality reduction technique which links words of original texts with the same context and discover themes that run through them, called topics. [Aggarwal and Zhai (2012)]
- **Definition 3: Document Clustering** is the segmentation of documents in a corpus into homogenous groups, using the words or terms within each document as features in the model. [Manning and Hinrich (1999)]

Model Building Procedure

To determine the demand and industry-required competencies for the four (4) data science and analytics roles, the study performed the following tasks were performed:

- (1) Skills sets or competencies were extracted from the descriptions of online job posts.
- (2) Extracted jobs were categorized into groups with similar skills requirements or role responsibilities.

Online Job Post Scraping

Job posts from May 1, 2019 to July 21, 2019 were scraped from Jobstreet.com.ph, using keywords that are related to the data profession (e.g. roles, tools, and APEC DSA competencies). Specifically, the lists of keywords are given in Table A.1, A.2, and A.3:

Table A.1 Keywords per APEC Data Science and Analytics Competency

Competency	Keyword
Domain Knowledge and Application	domain
Domain Knowledge and Application	industry
Domain Knowledge and Application	business
Domain Knowledge and Application	practice
Domain Knowledge and Application	procedure
Data Management and Governance	governance
Data Management and Governance	data security
Data Management and Governance	information security
Data Management and Governance	policy
Data Management and Governance	privacy
Data Management and Governance	MDM
Operational Analytics	analysis
Operational Analytics	process
Operational Analytics	performance
Operational Analytics	operations
Operational Analytics	metrics
Operational Analytics	six sigma
Operational Analytics	business Intelligence
Operational Analytics	MIS
Data Visualization and Presentation	visualization
Data Visualization and Presentation	data storytelling
Data Visualization and Presentation	graph
Data Visualization and Presentation	charts
Data Visualization and Presentation	insights
Data Visualization and Presentation	report
Data Visualization and Presentation	dashboard
Research Methods	research
Research Methods	hypothesis
Research Methods	experiment
Data Engineering Principles	ETL
Data Engineering Principles	data warehouse
Data Engineering Principles	big data
Data Engineering Principles	data lake
Data Engineering Principles	architecture
Statistical Techniques	statistics
Statistical Techniques	quantitative
Statistical Techniques	qualitative
Statistical Techniques	regression
Statistical Techniques	ANOVA
Data Analytics Methods and Algorithms	algorithm
Data Analytics Methods and Algorithms	machine learning
Data Analytics Methods and Algorithms	statistical
Data Analytics Methods and Algorithms	mathematical
Data Analytics Methods and Algorithms	predictive
Data Analytics Methods and Algorithms	data mining
Data Analytics Methods and Algorithms	econometrics
Computing	SQL
Computing	programming
Computing	software
Computing	hardware
Computing	computing
Computing	IoT
Computing	cloud

Source: Authors' list based on APEC's definitions of the 10 Data Science and Analytics Competencies

Table A.2: Keywords per Data Science and Analytics Role

Job Role	Keyword
Data Steward	steward
Data Steward	governance
Data Steward	privacy
Data Steward	policy
Data Steward	data security
Data Engineer	ETL
Data Engineer	data warehouse
Data Engineer	big data
Data Engineer	database
Data Engineer	data engineer
Data Engineer	programmer
Data Engineer	developer
Data Engineer	solution engineer
Data Engineer	solution architect
Data Scientist	statistic
Data Scientist	math
Data Scientist	statistical
Data Scientist	mathematical
Data Scientist	predictive
Data Scientist	algorithm
Data Scientist	data science
Data Scientist	actuarial
Data Scientist	statistician
Functional Analyst	industry
Functional Analyst	domain
Functional Analyst	operations
Functional Analyst	business
Functional Analyst	analyst

Source: Authors' list based on APEC's definitions of the 10 Data Science and Analytics Competencies

Table A.3: Keywords for Other Categories

Category	Keyword
Tools	Python
Tools	SAS
Tools	SPSS
Tools	RapidMiner
Tools	Knime
Tools	R
Tools	Eviews
Tools	Minitab
Tools	Informatica
Tools	Oracle
Tools	Teradata
Tools	Tableau
Tools	Power BI
Tools	Stata
Tools	Excel
Tools	Access
Tools	Qlik
Tools	Matlab
Tools	VBA
Others	analytics
Others	data

Source: Authors' list

Overall, 150,869 unique posts were scraped. The fields extracted include:

Table B: List of Fields

Column Name	Description
job_post_id	Job Post ID
job_title	Job Title
company_name	Company Name
job_salary	Job Salary
yrs_exp	Years of Experience
job_location	Job Location
avg_processing_time	Average Processing Time
company_industry	Company Industry
company_size	Company Size
job_desc	Job Description
job_post_dt	Job Post Dt
job_close_dt	Job_Close Dt

Source: Authors' list

For skills extraction and job segmentation, job description will mainly be used.

For faster data discovery, a random sample of size 20,574 (13.6%) was selected from the 150,869 unique posts. The algorithms described were performed on this representative sample.

Skills Pattern Matching and Tokenization

Based on a similar study done by Wang (2017), there is a need for a supervised approach in extraction of features (e.g. skills, courses, responsibilities, tools) from the job descriptions. Most descriptions have a lot of noise, due to inclusion of equal opportunity statements, application instructions, and company benefits. In an effort to remove the effect of these noises on the results of subsequent analyses, a predetermined set of skills from Wang (2017) was used as vocabulary (see Appendix B). From the job description column, another field for skills set was created, which only retained words or phrases that match the items in the predetermined skills list.

Figure A. Illustration of Skills Pattern Matching

job_post_id	job_desc	skills_set
8767788	Strategic Analytics is a new, growing team at Arch. The team develops innovative predictive models and analytical tools to improve profitability, growth, and operational efficiency . This position will help support predictive modeling needs across a growing portfolio of high-profile advanced analytics projects with the Strategic Analytics team. As a key member of the team you will also play a role in advancing predictive modeling capabilities, enabling Arch to make better decisions. REQUIREMENTS: Predictive modeling experience in a professional setting, using tools such as R , SQL , Python or SASWilling to work on Mid Shift (3pm-12am Mon-Fri)	['Predictive modeling', 'SQL', 'R', 'advanced analytics', 'analytical', 'analytics', 'Analytics', 'Python', 'operational efficiency', 'predictive modeling', 'modeling']

Source: Results from Authors' online job scraping

Unique skills were extracted from the skills sets (i.e. tokenization). For easy classification of these skills into similar groups based on the pattern by which they appear in the job posts, their term frequency-inverse document frequency (TF-IDF) statistics were calculated. TF-IDF measures the importance of a skill to a job post in a collection of job posts. It consists of 2 components [Weiss et. al (2015)]:

(1) *Term frequency (TF)* is defined as:

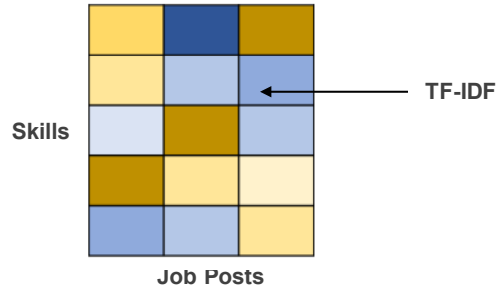
$$TF_{\text{skill}} = \frac{\text{No. of times the skill appears in a job post}}{\text{Total no. of skills in a job post}}$$

(2) *Inverse Document Frequency* is a scaling factor computed as:

$$IDF_{\text{skill}} = \log \frac{\text{Total no. of job posts}}{\text{No. of job posts with the skill in it}}$$

Higher values for TF indicate that the skill appears more frequently in the job post. Lower values for IDF indicate that the skill is common across job posts. Hence, the higher the TF-IDF value, the more the skill can distinguish the job posts from each other. The structure of the table that were used in the succeeding analyses is given below:

Figure B. Skill-Job Post Matrix



Source: Results from Authors' online job scraping

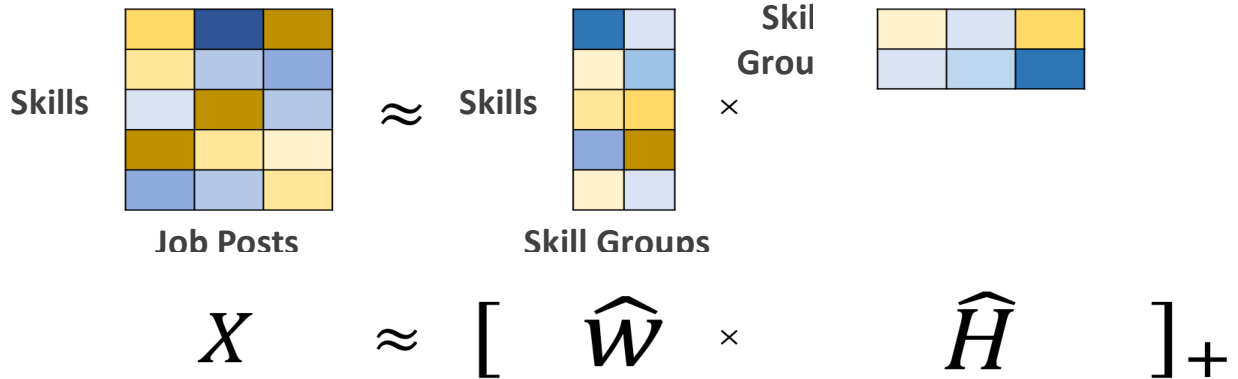
Topic Modeling

The 4,219 extracted skills still needed to be categorized into more interpretable groups or sets. This was done by topic modeling, where the underlying themes unifying the different skills are explored.

There are two classifications of topic modeling techniques—probabilistic and non-probabilistic. The most popular method under each is latent Dirichlet allocation (LDA) and non-negative matrix factorization (NMF), respectively [Cheng et. al (2013)]. The latter was employed in the analysis, because it works better and faster than the former when learning keywords or short texts [Cheng et. al (2013), Chen et. al (2016)].

NMF is a matrix factorization method that decomposes the skill-job post matrix $X_{(m \times n)}$, into 2 matrices, $\hat{W}_{(m \times k)}$ and $\hat{H}_{(k \times n)}$, where m = number of skills (4,219), n = number of job descriptions (20,574), and k = number of skills group (refer to Figure C) [Chen et. al (2016)]. The matrix \hat{H} contains the mapping of skills groups to the different job posts.

Figure C. NMF Algorithm



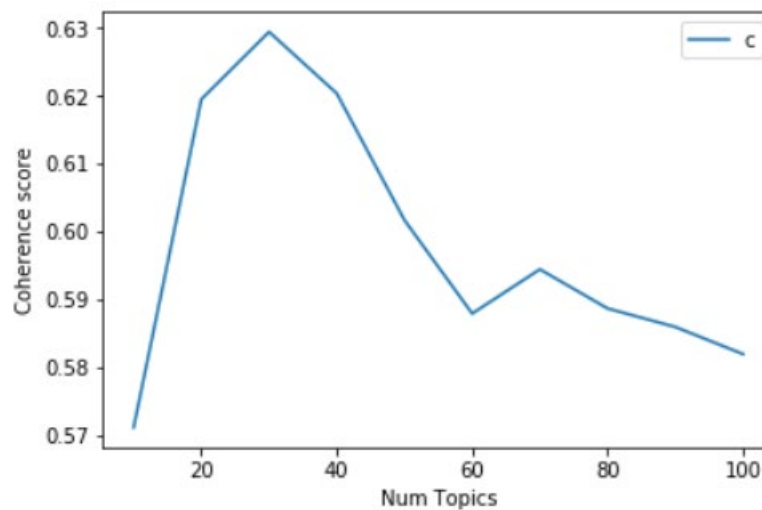
Source: Results from Authors' online job scraping

Since NMF will not automatically provide the optimal number of skills groups, the selection was based on topic coherence score. Coherence score measures whether pairs of top skills in a skills group tend to co-occur together relative to the whole set of job posts. Several iterations of NMF, varying the number of skills groups from 10 to 100, were ran. For each iteration, the coherence score was computed per skill group, then aggregated. The iteration with the highest aggregated score determines the optimal number of groups.

Röder et. al (2015) reviewed different variants of coherence measures and found that C_v is the one most aligned with human interpretability. Hence, this was the score method used.

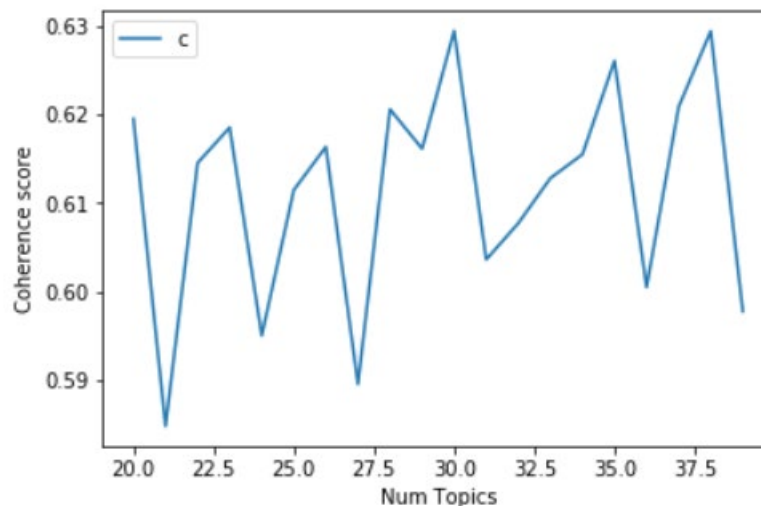
Results suggest that the topic coherence is highest when the number of skills groups chosen is 30.

Figure D.1: Topic Coherence Score for Number of Topics 10 to 100 by 10



Source: Results from Authors' online job scraping

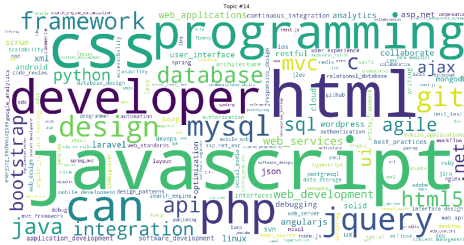
Figure D.2: Topic Coherence Score for Number of Topics 20 to 40 by 1



Source: Results from Authors' online job scraping

For a better understanding of what the skills groups represent, they were visualized using word clouds. The size of the words or phrases shows the weight of the skills in characterizing the group. Refer to Figure E for a sample and to Appendix B for all the complete set of skills groups word clouds.

Figure E: Sample Skills Groups Word Clouds



Skill Group 14



Skill Group 15

Source: Results from Authors' online job scraping

The skills group-job post matrix (recall matrix \hat{H}) resulting from the NMF procedure shows the percentage of each skills group in the requirements of job positions. The sum of these weights is equal to 100%.

Figure F. Sample Skills Group Weights by Job Post

job_post_id	Skills_Group_0	Skills_Group_1	Skills_Group_2	Skills_Group_3	Skills_Group_4	Skills_Group_5	Skills_Group_6	Skills_Group_7	Skills_Group_8	Skills_Group_9	Skills_Group_10	Skills_Group_11	Skills_Group_12	Skills_Group_13	Skills_Group_14	Skills_Group_15	Skills_Group_16	Skills_Group_17	Skills_Group_18	Skills_Group_19	Skills_Group_20	Skills_Group_21	Skills_Group_22	Skills_Group_23	Skills_Group_24	Skills_Group_25	Skills_Group_26	Skills_Group_27	Skills_Group_28	Skills_Group_29	TOTAL
8782678	0.0%	9.8%	0.0%	3.2%	0.0%	18.1%	0.0%	4.3%	0.0%	0.0%	0.0%	0.0%	2.3%	11.1%	0.0%	0.0%	0.0%	0.0%	4.0%	22.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	4.6%	19.9%	0.0%	0.0%	100.0%
8781741	0.0%	0.0%	0.0%	0.0%	0.0%	3.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	4.1%	0.0%	0.0%	0.0%	0.0%	6.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.5%	0.0%	54.1%	0.0%	29.9%	100.0%
8781262	0.0%	0.0%	0.0%	0.0%	9.3%	0.0%	0.0%	0.0%	0.0%	4.1%	0.0%	0.0%	0.0%	0.0%	0.0%	6.7%	0.0%	0.0%	0.0%	5.5%	0.0%	0.0%	14.8%	0.0%	0.0%	10.1%	0.0%	49.5%	0.0%	0.0%	100.0%
8779097	0.0%	0.0%	0.0%	0.0%	2.8%	0.0%	0.0%	0.0%	5.2%	0.0%	0.0%	0.0%	0.0%	0.0%	13.2%	1.4%	7.8%	0.0%	0.0%	12.3%	0.0%	4.3%	17.5%	0.0%	0.0%	0.0%	0.0%	33.5%	0.0%	1.9%	100.0%
8776774	24.7%	0.0%	0.0%	0.0%	14.3%	0.0%	0.0%	0.0%	4.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	15.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	41.5%	0.0%	0.0%	100.0%

Source: Results from Authors' online job scraping

These 30 skills group are then mapped to the 10 APEC Data Science and Analytics (DSA) competencies.

Table C. Skills Group–Competency Mapping

DSA Competency	Skills Group No.
Domain Knowledge	0, 12, 21
Data Governance	6, 11, 13, 16
Operational Analytics	2, 8, 20
Data Visualization	17
Data Engineering	3, 4, 7, 14, 25, 27, 29
Statistical Techniques	23
Methods & Algorithms	15
Computing	1, 22, 24, 26
21 st Century Skills	5, 9, 10, 18, 19, 28

Source: Results from Authors' online job scraping

To get the competency weights to be obtained per job post, the weights of the skills groups mapped to the same competency are combined (refer to Figure G for illustration).

Figure G. Illustration of Derivation of Competency Weights

Job_post_id	Skills_Group_21	Skills_Group_22	Comp_Domain_Knowledge
8782678	0.0%	0.0%	0.0%
8781741	0.0%	0.0%	0.0%
8781262	0.0%	14.8%	14.8%
8779097	4.3%	17.5%	21.8%
8776774	0.0%	0.0%	0.0%

Job_post_id	Comp_Domain_Knowledge	Comp_Data_Governance	Comp_Ops_Analytics	Comp_Data_Viz	Comp_Data_Engg	Comp_Stat_Techniques	Comp_Algorithms	Comp_Computing	Comp_21stCentury_Skills	TOTAL
8782678	0.0%	11.1%	44.9%	9.8%	4.6%	4.3%	19.9%	3.2%	2.3%	100.0%
8781741	0.0%	5.6%	10.4%	0.0%	29.9%	0.0%	54.1%	0.0%	0.0%	100.0%
8781262	14.8%	10.1%	14.8%	0.0%	0.0%	6.7%	49.5%	4.1%	0.0%	100.0%
8779097	21.8%	0.0%	28.3%	0.0%	7.1%	9.2%	33.5%	0.0%	0.0%	100.0%
8776774	0.0%	0.0%	14.3%	24.7%	4.5%	15.0%	41.5%	0.0%	0.0%	100.0%

Source: Results from Authors' online job scraping

As in the skills group level, the sum of the competency weights should also equal to 100%. The weights now represent the contribution of the competencies to the requirements of the job positions.

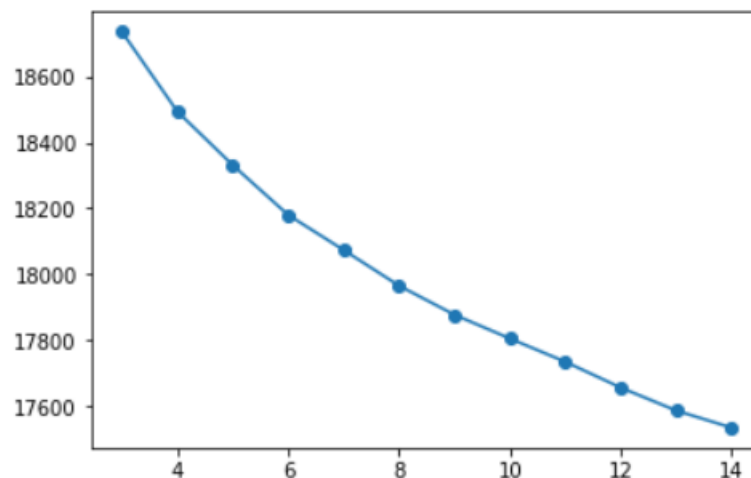
Document Clustering

Instead of grouping skills, the job posts were grouped in order to find among them the DSA roles—data steward, data engineer, data scientist, functional analyst, and analytics manager. Grouping was done by applying clustering analysis on the skill-job post TF-IDF matrix.

There are several algorithms to implement clustering [Xu (2005)], but k-means method was the one used in the analysis.

As with NMF, k-means algorithm does not automatically suggest the best number of clusters. Hence, the number of groups from 2 to 15 was tested. Inertia, which is the within-cluster sum of squares, was obtained from each iteration. Plotting the clustering errors against the number of clusters, the point at which the improvement starts to become marginal is ideally chosen as the "*elbow criterion*" (stopping criterion) [Arthur and Vassilvitskii (2007)].

Figure H. Inertia by Number of Clusters



Source: Results from Authors' online job scraping

Since the results did not show a clear elbow, iterations with the number of groups from 4 to 10 were revisited (close to the number of data roles the study wants to discover). The number of job groups was subjectively selected such that the top skills per group are as related as possible and the skills across job groups are as different as possible. The differentiation of the job titles across job groups were also considered in the selection of the number of groups. Ten job groups were generated.

Figure I. Top Terms of the Final Set of Clusters



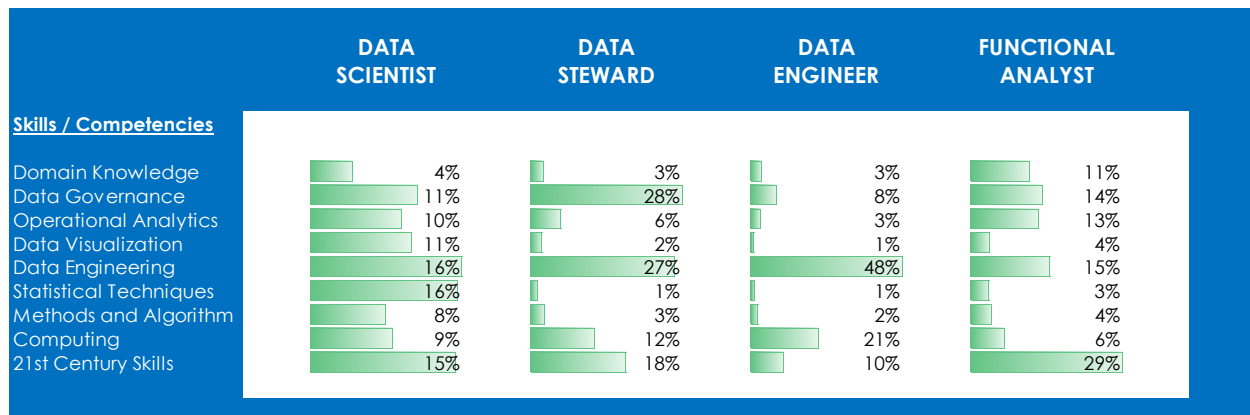
Source: Results from Authors' online job scraping

These clusters were further refined, following the below reclassification:

New Groupings	Original Clusters
Data Scientist	Data Scientist
Licensed Manager / Analyst	Dropped
Data Steward	Data Steward
Data Engineer	Data Engineer, Software / App Developer
Functional Analyst	Audit Manager / Analyst, Finance Analyst, Inventory / Warehouse Manager, Tech Support Specialist, Analytics Manager

Source: Results from Authors' online job scraping

Figure J. Employers' DSA Role Profiles



Source: Results from Authors' online job scraping

Appendix B. List of pre-determined set of skills by Wang (2017)



vocab.txt

Appendix C. Distribution table of DSA employees' degree programs

	DSA Degree Program	COUNT	% Count
1.	Computer Science	11	10%
2.	Business Administration	11	10%
3.	Statistics	10	9%
4.	Applied Math/Math	9	8%
5.	Information Technology	8	7%
6.	Information Science	7	6%
7.	Economics	6	5%
8.	Physics/Applied Physics	6	5%
9.	Industrial Engineering	4	4%
10.	Civil Engineering	4	4%
11.	Electrical Engineering	3	3%
12.	Electronics and Communications Engineering	3	3%
13.	Accountancy	3	3%
14.	Geodetic Engineering	3	3%
15.	Communications-related	3	3%
16.	Chemistry	2	2%
17.	Humanities	2	2%
18.	Political Science	2	2%
19.	Psychology	2	2%
20.	Biology	1	1%
21.	Commerce	1	1%
22.	Computer Applications	1	1%
23.	Computer Engineering	1	1%
24.	Fine Arts	1	1%
25.	Legal Management	1	1%
26.	Mechanical Engineering	1	1%
27.	Public Health	1	1%
28.	Sociology	1	1%
29.	Tourism	1	1%
30.	Geography	1	1%
	TOTAL PARTICIPANTS	110	100%

Source: Survey Responses from Data Science and Analytics-related practitioners

Appendix D. Employment Share per Major Industry Group (2013-2017)

By Major Industry Group	Employment Share (in thousands)					Compound Annual Growth rate ²⁶
	2013	2014	2015	2016	2017	
Agriculture	11836	11801	11294	11064	10261	-3%
Industry	5936	6166	6275	7159	7371	4%
Services	20345	20682	21172	22775	22703	2%
Total	38, 117	38, 649	41, 228	40, 998	40, 335	

Source: Philippine Statistics Authority 2013-2017, Labor Force Survey as cited from *JobsFit 2022 Report by Bureau of Local Employment, Department of Labor and Employment; PSA CISLE Report, 2018*

Appendix E. Employment Size of Industry subclasses of BPM subsector

Section	Code	PSIC Industry sub-classes (non-BPM)	Employees			
			2013	2014	2015	2016
BPM (sections I and J)	J58190	Other publishing activities	492	466	349	394
I: Information and Communication	J58200	Software publishing	924	1,039	2,281	1,852
N: Admin Support and Services	J59110	Motion picture, video and television programme activities*	67	284	445	418
	J59120	Motion picture, video and television programme post-production activities*	935	846	939	708
	J62010	Computer programming activities	36,023	35,951	39,503	38,808
	J62020	Computer consultancy and computer facilities management activities	12,041	10,766	14,609	11,476
	J62090	Other information technology and computer service activities	3,833	4,372	4,049	3,780
	J63111	Data processing	19,163	18,212	17,684	14,052

²⁶ Computed by Authors based on standard Compound Annual Growth Rate (CAGR) formula

Section	Code	PSIC Industry sub-classes (non-BPM)	Employees			
			2013	2014	2015	2016
	J63112	Website hosting services	533	738	294	905
	J63113	Application hosting services	721	564	588	448
	N78103	On-line employment placement agencies	106	104	73	78
	N82211	Customer relationship management activities	298,485	354,140	390,799	420,763
	N82212	Sales and marketing (including telemarketing) activities	52,940	72,539	72,546	80,442
	N82219	Other call centers activities (voice), n.e.c.	7,366	11,070	9,265	9,512
	N82221	Finance and accounting activities	6,948	6,699	6,809	7,227
	N82222	Human resources and training activities	404	512	2,742	2,434
	N82223	Administrative support activities	591	443	1,484	1,691
	N82224	Document processes activities	189	119	173	928
	N82225	Payroll maintenance and other transaction processing activities	39	42	297	700
	N82226	Medical transcription activities	2,066	2,058	1,886	2,945
	N82227	Legal services activities	s	s	s	s
	N82228	Supply chain management activities	s		s	s
	N82229	Other non-voice related activities, n.e.c.	2,008	2,290	2,850	6,627
	N82291	Engineering outsourcing activities	806	937	822	1,562
	N82292	Product development activities	s	s		

Section	Code	PSIC Industry sub-classes (non-BPM)	Employees			
			2013	2014	2015	2016
	N82293	Publishing outsourcing activities	2,802	596	824	495
	N82294	Research and analysis activities	332	162	183	291
	N82295	Intellectual property research and documentation activities		s	s	144
	N82296	Security outsourcing activities	91		s	
	N82299	Other back office operations activities, n.e.c	nad	823	1,618	1,141
		TOTAL	449,905	525,772	573,112	609,821

Source: Philippine Statistics Authority 2013-2017, Labor Force Survey as cited from *JobsFit 2022 Report by Bureau of Local Employment, Department of Labor and Employment*; PSA CISLE Report, 2018

Appendix F. Employment size of Analytics-related activities from PSIC's list and definition of industries (2013-2016)

Industry (based on PSIC)	Code	Industry sub-classes	2013		2014		2015		2016	
			Employees	%	Employees	%	Employees	%	Employees	%
C: Manufacturing	C26200	Manufacture of computers and peripheral equipment and accessories	55,288	35.77%	60,096	34.41%	73,462	36.80%	58,224	33.52%
K: Financial and Insurance Services	K66290	Other activities auxiliary to insurance and pension funding	aggregated	NA	aggregated	NA	255	0.13%	191	0.11%
M: Professional, Scientific and Technical Activities	M70200	Management consultancy activities	13,729	8.88%	24,795	14.20%	26,862	13.45%	27,396	15.77%
	M71109	Other technical activities related to architectural and engineering	nad	NA	nad	NA	2,530	1.27%	2,902	1.67%
	M72101	Research and experimental development	1,413	0.91%	1,557	0.89%	610	0.31%	723	0.42%

Industry (based on PSIC)	Code	Industry sub- classes	2013		2014		2015		2016	
			Employees	%	Employees	%	Employees	%	Employees	%
		in natural sciences								
	M72102	Research and experimental development in engineering and technology	nad	NA	nad	NA	688	0.34%	342	0.20%
	M72103	Research and experimental development in health sciences	nad	NA	nad	NA	497	0.25%	465	0.27%
	M72104	Research and experimental development in agricultural sciences	nad	NA	nad	NA	245	0.12%	257	0.15%
	M72200	Research and experimental development on social sciences and humanities	248	0.16%	886	0.51%	531	0.27%	1,034	0.60%
	M72300	Research and experimental development in information technology	2586	1.67%	2,604	1.49%	3,492	1.75%	3,686	2.12%
	M72400	Research and experimental development services, n.e.c.	nad	NA	56	0.03%	18	0.01%	s	NA
	M73200	Market research and public opinion polling	6037	3.91%	8,648	4.95%	6,728	3.37%	5,713	3.29%
	M749	Other professional, scientific and technical activities, n.e.c.	1703	1.10%	3,400	1.95%	2,921	1.46%	nad	NA
J: Information and Communication	J58200	Software publishing	924	0.60%	1,039	0.59%	2,281	1.14%	1,852	1.07%
	J62010	Computer programming activities	36,023	23.30%	35,951	20.58%	39,503	19.79%	38,808	22.34%
	J62020	Computer consultancy and computer facilities management activities	12,041	7.79%	10,766	6.16%	14,609	7.32%	11,476	6.61%

Industry (based on PSIC)	Code	Industry sub- classes	2013		2014		2015		2016	
			Employees	%	Employees	%	Employees	%	Employees	%
	J62090	Other information technology and computer service activities	3,833	2.48%	4,372	2.50%	4,049	2.03%	3,780	2.18%
	J63111	Data processing	19,163	12.40%	18,212	10.43%	17,684	8.86%	14,052	8.09%
	J63112	Website hosting services	533	0.34%	738	0.42%	294	0.15%	905	0.52%
	J63113	Application hosting services	721	0.47%	564	0.32%	588	0.29%	448	0.26%
N: Administrative and Support Services	N82294	Research and analysis activities	332	0.21%	162	0.09%	183	0.09%	291	0.17%
	N82299	Other back office operations activities, n.e.c	nad	NA	823	0.47%	1,618	0.81%	1,141	0.66%
		TOTAL	154,574	100%	174,669	100.00%	199,648	100.00%	173,686	100.00%

Source: Philippine Statistics Authority. (2013-2016). *Annual Survey of Philippine Business and Industry (ASPBI)*.

Retrieved from <https://psa.gov.ph/statistics/survey/business-and-industry/index>