

Navigating the Digital Frontier: A Conceptual Framework for Data Science and Analytics Integration in Tourism Education and Research

Sahil Sharma

sahilsharmahimalaya@gmail.com

Himachal Pradesh University

Roman Egger

MODUL University Vienna

Research Article

Keywords: Machine Learning, Curriculum Design, Digital Skills Gap, Hospitality Management, Job Advertisement, Grounded Theory

Posted Date: October 20th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7874079/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

The growing importance of big data and data analytics has made these skills indispensable for the tourism industry, necessitating an urgent revision of tourism education curricula. Previous research, however, lacks both ontological clarity and empirical findings regarding the specific skills and competencies required. To address this knowledge gap, this study employs a grounded theory approach through the analysis of job advertisements. The results first confirm a significant industry demand for data science and analytics (DSA) roles, answering why this topic is critical. Second, the study establishes an ontology of DSA by conceptualizing and categorizing the required skills and competencies, thus defining what is needed. Finally, by identifying specific teams and job focus areas, the research clarifies where and how these skills can be applied in practice and integrated into educational syllabi. The study concludes by discussing key theoretical, methodological, and practical implications that underscore its contribution to the field.

1. Introduction

Rapid technological advancements are reshaping the traditional disciplinary boundaries within tourism, prompting a critical reassessment and expansion of existing theories and frameworks (Sigala, 2018). As digital transformation accelerates, advanced analytical capabilities such as Data Science and Analytics (DSA)—previously viewed as peripheral or supplementary to tourism education—are now becoming fundamental requirements for professional competence within the sector (Buhalis et al., 2019; Carlisle et al., 2023; Sigala, 2018). Consequently, tourism scholars advocate for an immediate and comprehensive restructuring of tourism curricula, recommending a broader integration that includes humanities, science, and engineering principles, particularly emphasizing DSA skills (Hsu et al., 2017).

The impetus for incorporating DSA into tourism education stems significantly from the proliferation of big data and the current skills gap among tourism graduates and professionals (Sharma, 2023, 2024; Sigala, 2018). Hsu (2018, p. 2) explicitly highlights this gap, stating: “A tourism data science program will allow students to develop skills necessary to analyze and interpret big data, as well as learn the design tools and methodologies to obtain such data.”

Research examining digital competencies among tourism professionals has underscored the increasing relevance of DSA skills, although much of this work has not directly addressed curriculum design and educational planning (Carlisle et al., 2023; Lamest & Brady, 2019; Minor et al., 2024). Concurrently, some studies specifically addressing educational planning for tourism have emerged (Adeyinka-Ojo et al., 2020; Busulwa et al., 2022; Ogbeide et al., 2021), yet these investigations exhibit two critical limitations. First, existing research predominantly comprises conceptual analyses derived solely from literature reviews, lacking direct empirical insights into actual industry demands. Second, discussions regarding DSA have remained largely theoretical and abstract, failing to offer granular, detailed guidance regarding specific skills, techniques, or conceptual competencies required by the tourism industry. These gaps

limit a precise understanding of DSA among tourism educators and policymakers, complicating effective integration into curricula and professional training programs.

In response to these research limitations, the present study focuses explicitly on DSA within the broader spectrum of digital skills in tourism, pursuing two primary objectives:

1. To empirically investigate the specific job roles, requisite skills, and professional competencies related to DSA in the tourism industry.
2. To conceptualize and structure the key skills, techniques, and concepts currently in demand, thereby supporting precise and effective tourism education planning.

Methodologically, this study employs an integrated approach combining Job Advertisement Analysis (JAA) with Grounded Theory. This methodological innovation enhances traditional JAA approaches, providing robust empirical validation directly aligned with current industry demands. Moreover, it addresses a historical disconnect between academia and industry, an issue extensively documented by scholars (Airey & Nightingale, 1981; Hsu et al., 2017; Tribe, 2002). The detailed methodological description included in this paper underscores the rigor and replicability of the proposed research framework.

The significance of this research is underscored by the rapid and sustained growth of DSA roles within major tourism enterprises, including Expedia, TripAdvisor, Airbnb, and Make My Trip, all of which have established specialized departments dedicated to data science and analytics. Reflecting this industry shift, pioneering educational institutions worldwide are integrating DSA into tourism curricula (see **Appendix A**). This growing trend indicates a clear trajectory toward broader global adoption, highlighting the critical importance of informed curriculum planning grounded in empirical research.

Thus, this study not only bridges the existing gap between theory and practice but also provides concrete insights and strategic direction for embedding essential DSA competencies within tourism education, thereby aligning academic offerings with industry demands and enhancing graduate employability in an increasingly data-driven tourism landscape.

2. Literature Review

2.1 Applications of Data Science and Analytics (DSA) in Tourism Research

Despite the limited formal integration of data science and artificial intelligence methods within traditional tourism education, the application of Data Science and Analytics (DSA) methodologies has notably expanded in tourism research over recent years (Fuchs et al., 2014; Mariani et al., 2018). Much of this scholarly activity has leveraged big data, predominantly sourced from user-generated content (UGC) on social media platforms (Li et al., 2018; Mariani, 2020). However, the use of device-generated and

transactional big data remains relatively underdeveloped, largely constrained by issues of data accessibility, privacy concerns, and security barriers.

Initially, tourism research adopting data science techniques primarily relied on text-based analysis methodologies (Xiang et al., 2015, 2017). However, recent literature indicates an emerging focus on image and video analytics, underscoring a more diversified methodological landscape (Barnes & Kirshner, 2021; Zhan et al., 2024; Zhu et al., 2024). Additionally, advanced machine learning and deep learning techniques—such as Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Support Vector Machines (SVM), Back Propagation Neural Networks (BPNN), Artificial Neural Networks (ANN), Graph Neural Networks (GNN), and Autoregressive Integrated Moving Average (ARIMA)—are increasingly common in contemporary tourism analyses (Doborjeh et al., 2022). The application areas for these methodologies are diverse, encompassing consumer behavior analysis, marketing management, and demand forecasting (Lv et al., 2022). This evolving research landscape highlights the necessity of incorporating robust DSA training into tourism education to adequately prepare future researchers and professionals for the industry's data-driven transformation.

2.2 Demand for Incorporating Data Science and Analytics (DSA) into Tourism Education

Parallel to research developments, there is a growing consensus on the urgent need to integrate DSA competencies into tourism education, driven primarily by the rapid digitalization of the industry and the evident digital skills gap among tourism professionals and graduates (Busulwa et al., 2022; Carlisle et al., 2023; Lamest & Brady, 2019; Minor et al., 2024). For instance, Busulwa et al. (2022) underscore the disruptive influence of digital transformation on data availability, explicitly identifying data analytics and management competencies as critical for contemporary hospitality managers. Carlisle et al. (2023) similarly identify substantial deficiencies in data analytics skills within the European tourism workforce, pointing toward a paradigm shift where skills traditionally aligned with technology-focused roles are increasingly essential for tourism professionals (Buhalis et al., 2019; Buhalis & Law, 2008; Carlisle et al., 2023).

Furthermore, empirical evidence highlights the industry-wide necessity of digital competencies. Lamest and Brady (2019) illustrate the relevance of big data technologies within hospitality management, advocating that managers must acquire appropriate analytics skills. In a similar vein, Minor et al. (2024) reveal significant digital skills gaps across the United Kingdom's tourism sector, reinforcing the argument for curricular integration of digital and analytical skills.

From the educational perspective, scholars argue for an immediate curricular transformation that extends beyond the traditional confines of business and social sciences, integrating scientific and engineering principles (Hsu et al., 2017). Hsu (2018) specifically advocates for dedicated "Tourism Data Science" programs. Supporting this perspective, Ogbeide et al. (2021) demonstrate that curricula integrating DSA competencies correlate positively with enhanced employment prospects and higher salaries for tourism graduates. Recent initiatives such as the sports analytics curriculum at Purdue

University (Ogbeide et al., 2021) and the Industry 4.0 program at the University of Calabria (Bilotta et al., 2021) exemplify successful practical implementations of DSA competencies within tourism education.

Forecasting future industry demands, Ivanov (2020) emphasizes the inevitability of a growing need for tourism professionals proficient in data science and analytics—not necessarily to create technological applications themselves, but to effectively understand and leverage their implications for organizational and operational efficiency.

2.3 Understanding Industrial Needs in Tourism Education Research

A persistent issue within tourism education research is the historical disconnect between academic curriculum development and actual industry requirements. Initial explorations by Airey and Nightingale (1981) marked early attempts to align tourism education more closely with industry demands by examining occupational profiles and requisite knowledge. Despite these foundational studies, subsequent scholarship consistently highlights a persistent gap in academia-industry collaboration, underscoring inadequate industrial input into curriculum design (Elnasr Sobaih & Jones, 2015; McCartney & Kwok, 2023; Mohanty & Mohanty, 2019; Tribe, 2002).

More recent analyses reiterate the sustained low involvement of industry in academic program design, despite long-standing recognition of this issue (Hsu et al., 2017). This ongoing disconnect contributes significantly to the lag in embedding critical DSA skills within tourism curricula, despite clear industry demands for such competencies evidenced by the establishment of specialized DSA departments in major companies.

Addressing this gap, the present study employs a novel methodological approach—Job Advertisement Analysis (JAA)—enhanced through Grounded Theory to systematically capture and analyze specific industrial demands for DSA skills. This innovative approach not only provides empirical insights into industry requirements but also directly informs tourism curriculum planners and educators, facilitating precise alignment between educational offerings and industry expectations. Consequently, this methodological innovation represents a valuable step toward bridging the persistent academia-industry divide, thus enhancing the relevance and responsiveness of tourism education programs.

3. Research Methodology

3.1 Job Advertisement Analysis using Grounded Theory

Job Advertisement Analysis (JAA) leverages publicly available job advertisement data for diverse research applications. Due to its accessibility and abundance, JAA has proven highly effective across various disciplines, including library sciences and medical sciences (Choi & Rasmussen, 2009; Meyer, 2019). Within tourism research, JAA has historically offered valuable insights, though recent studies highlight a notable underutilization of this approach, suggesting it has significant untapped potential for

informing tourism education (Arcodia & Barker, 2002; Bernstein & Woosnam, 2019; Robinson et al., 2010; Arcodia et al., 2020).

Traditionally, JAA studies have employed content analysis for processing and interpreting job advertisements. However, this research adopts Grounded Theory (GT) to provide a more robust theoretical grounding for the empirical insights derived from job advertisements. GT offers distinct advantages over content analysis, notably its inductive coding processes, which facilitate deeper exploration of patterns within the data. Importantly, GT does not constrain data to predefined coding frameworks, enabling researchers to capture emerging concepts and relationships organically (Glaser, 1992; Glaser & Strauss, 1967). Specifically, this study uses Constructivist Grounded Theory, emphasizing the researcher's active role in interpreting data and co-constructing knowledge, aligning closely with pragmatist philosophical underpinnings (Charmaz, 2006). This methodological shift enriches the analysis of nuanced DSA roles within tourism, promoting a more comprehensive understanding of the industry's evolving skill demands.

Given the innovative integration of GT into JAA, the following sections detail the methodological procedures and analytical steps employed, facilitating replication and application in future tourism research.

3.2 Data Collection and Analysis: Application of Constructivist Grounded Theory

Grounded Theory encompasses iterative analytical processes such as memo-writing, constant comparative analysis, systematic coding, and theoretical sampling. Memo-writing throughout the research facilitated ongoing exploration of ideas, enabling constant comparative analysis across data points, codes, concepts, and categories, ultimately guiding theoretical development. Theoretical sampling, another key GT method, involved selectively gathering additional data to refine and elaborate upon emerging theoretical categories until conceptual saturation was achieved (Charmaz, 2006).

Data Collection

In Constructivist GT, data can originate from two categories of textual sources: extant texts, such as organizational reports, and elicited texts, such as interview transcripts. Extant texts (ETs) provide an unobtrusive, readily available, and seemingly objective source of information suitable for repurposing to answer specific research questions (Charmaz, 2006). In this study, job advertisements served as ETs collected over a five-month period from December 2023 to April 2024, in 15-day intervals. Searches employed job aggregators including Indeed, Huntr.co, and Hired.com, using keywords such as “data science” or “analytics” combined with “tourism,” “travel,” or “hospitality.” Subsequent monitoring of key recruiter websites supplemented these aggregator results. The final dataset comprised 462 initial documents.

Contextual Positioning and Data Selection

Effective utilization of ETs within GT requires detailed contextual understanding (Ralph et al., 2014). This study recorded variables such as “Advertisement ID,” “Job Title,” “Location,” and “Company Name” using MaxQDA software, facilitating contextual positioning and rigorous data selection. The selection process, enhanced by exporting document variables to the R programming environment for duplicate detection, resulted in the removal of irrelevant, duplicated, temporary, or internship postings. After filtering, 346 high-quality job advertisements were retained for detailed analysis.

Coding Process

The coding process involved three iterative phases: initial, focused, and axial coding (Fig. 1). Initially, broad textual segments related to skills or job descriptions were identified, with additional attributes (degree and experience requirements) recorded separately. Categories such as “Job Level” and “Hiring Organization Type” emerged during this phase to capture higher-level conceptual distinctions.

In the focused coding phase, skill-related textual data underwent detailed word-by-word and line-by-line analysis, employing continuous comparative methods to identify and refine a heuristic coding structure. Given the diverse terminologies used by employers, considerable interpretive judgment was necessary. Team names associated with job roles were identified or inferred and recorded.

Axial coding elevated conceptual sub-categories into broader theoretical categories, refining skills into major categories and teams into Job Focus Areas (JFAs). Continuous comparative analysis across documents and emerging codes resolved ambiguities, particularly in cases of overlapping team roles across multiple JFAs. To maintain analytical coherence and conciseness, similar teams were merged when appropriate.

Theoretical Sampling and Reporting

Theoretical sampling, employed after initial category formation, involved further interrogation of the existing dataset to explore and substantiate relationships among categories and concepts. The large dataset of 346 documents was especially conducive to theoretical sampling, supporting deeper theoretical insights rather than statistical generalization (Charmaz, 2006).

Reporting and interpreting findings incorporated the integration of MaxQDA and R programming tools. This integration enabled sophisticated filtering and visualization of trends and patterns, using R’s Tidyverse package for effective data manipulation and graphical representation (Wickham, 2016; Wickham et al., 2019). Thus, the methodological rigor and analytical depth provided by this GT-enhanced JAA approach significantly advanced understanding of the current and evolving DSA skill demands in tourism.

4. Results

4.1 The Jobs, Hiring Organisations, Educational Requirements, and other document-level variables

In terms of Data Science and Analytics (DSA) positions, this study identified 346 permanent roles within the global tourism industry. The United States dominated these roles, accounting for 46.2% (Table 1), possibly attributable to the headquarters of major tourism corporations being located there. Given the study's exclusive use of English-language job advertisements, the actual number of DSA jobs in non-English-speaking markets like China, Germany, and France is likely higher. Other notable English-speaking markets significantly recruiting DSA professionals included the United Kingdom, Netherlands, Thailand, and India.

Table 1
Document variable statistics (N = 346)

	<i>n</i>	<i>%</i>		<i>n</i>	<i>%</i>
<i>Job Levels</i>			<i>Hiring Organisation Type</i>		
Entry Level Jobs	188	54.3	Accommodation Provider	176	50.9
Senior Level Jobs	89	25.7	Online Travel Agency (OTA)	122	35.3
Manager or Director Level Jobs	50	14.5	Review Platform	22	6.4
Senior Manager Level Jobs	19	5.5	Transportation (Mediator, Airline, Cruise)	20	5.8
Total	346	100.0	Misc (GDS, Restaurant Booking, Travel Tech.)	6	1.7
Degree Required			Total	346	100.0
PhD or Masters	121	35.0	Job Focus Areas		
NA	106	30.6	Marketing	173	50.0
Master's degree	40	11.6	Data Science and Analytics	105	30.3
Bachelor's degree	29	8.4	Financial Operations	28	8.1
Bachelor's degree or higher	23	6.6	Cyber Security	22	6.4
Bachelors/Masters/PhD	21	6.1	Human Resource and Management	18	5.2
PhD	6	1.7	Total	346	100.0
Total	346	100.0	Locations		
Degree Preferred			United States (US)	160	46.2
NA	261	75.4	United Kingdom (UK)	44	12.7
PhD	49	14.2	NA	33	9.5
PhD or Masters	17	4.9	Netherlands	31	9.0
Master's degree	15	4.3	Thailand	20	5.8
Bachelors/Masters/PhD	3	0.9	India	13	3.8
Bachelor's degree	1	0.3	China	11	3.2
Total	346	100.0	Switzerland	11	3.2
Experience			Israel	6	1.7
4 years or higher	56	16.2	Portugal	6	1.7

	<i>n</i>	<i>%</i>		<i>n</i>	<i>%</i>
<i>Job Levels</i>	<i>Hiring Organisation Type</i>				
3 years or higher	55	15.9	Singapore	4	1.2
5 years or higher	45	13.0	France	2	0.6
Some undefined experience	44	12.7	Spain	2	0.6
2 years or higher	43	12.4	Canada	1	0.3
NA	25	7.2	Germany	1	0.3
6 years or higher	24	6.9	Mexico	1	0.3
8 years or higher	23	6.6	Total	346	100.0
7 years or higher	14	4.0			
10 years or higher	8	2.3			
9 years or higher	4	1.2			
1 year or higher	3	0.9			
0 year or higher	1	0.3			
15 years or above	1	0.3			
Total	346	100.0			
Note: The <i>n</i> symbol refers to the frequency and the <i>%</i> sign refers to the percentage.					

To establish theoretical orientation for this research, "Company Names" and "Job Titles" were elevated into conceptual categories. Twenty companies or corporate ventures were initially identified and subsequently categorized into five *Hiring Organisation Types*. Table 1 reveals that "Accommodation Providers" and "Online Travel Agencies" (OTAs) dominated recruitment, collectively holding 85–86% of market share. This finding has significant implications for educational planning, as tourism education focused on data science and analytics (DSA) should align with these major recruiters' specific requirements.

The nature of work within these organizations can be understood through their recruitment teams. Accommodation Providers primarily sought candidates for Product Development, Supply Analytics (Rooms and Properties), Marketing, and Payment teams, while OTAs focused on Data Science, Analytics, Automation, Supply, and Machine Learning positions. Review Platforms (Table 1) predominantly recruited for customer experience teams, representing another potential educational focus area.

Job titles, though diverse, clustered into distinct hierarchical levels, leading to the development of four *Job Levels* categories (Table 1). Entry-level positions dominated at 54.3%, encompassing roles like "Data

Scientist" and "Data Analyst," often modified with prefixes such as "Junior" or "Staff" and specific team designations. Senior-level positions ranked second, including titles like "Team Leader" and "Principal Data Scientist." Management positions, distinguished by their supervisory responsibilities and extensive experience requirements, comprised titles such as "Director," "Manager," "Senior Manager," and "Vice President."

Regarding *educational degree requirements*, Table 1 demonstrates that PhD or Master's degrees were most frequently required, with PhD being the preferred qualification. These degrees were predominantly sought from quantitative disciplines including STEM, Marketing, Data Science, and Economics. Notably, tourism-specific degrees were never requested despite all positions being within the tourism industry.

The data reveals that DSA-related positions primarily target candidates with master's degrees, while PhD qualifications provide competitive advantage at entry level and become essential for advancement to senior positions.

For *experience requirements*, distinct patterns emerged across job levels: entry-level positions required 2–4 years, senior-level roles demanded 4–6 years, and managerial positions sought 5–10 years of experience. Domain knowledge (see Fig. 4 up next) represents another dimension of experience, with "Relevant Work Experience" being most prevalent, followed by "E-commerce," "Marketing," and "Travel Industry Experience."

This creates an interesting paradox: while tourism degrees were never specified as requirements, travel industry experience was actively preferred, suggesting that practical sector knowledge holds greater value than formal tourism education in DSA recruitment.

4.2 DSA Skill Distributions, Concepts, and Potentials for Tourism Education Transformation

Understanding skill components provides the foundation for integrating DSA into tourism education. Figure 2 illustrates the distribution of the top 30 skills by frequency and tracks the seven most prominent skills across job levels. Four skills emerged as highly demanded: "Python," "SQL," "Communication & Data Storytelling," and "R."

Analysis revealed that employers expected expertise in either Python or R, with functional knowledge of the alternative language. While developing programming proficiency may traditionally challenge tourism education programs, large language models (LLMs) such as ChatGPT and GitHub Copilot have transformed programming instruction for management students from a significant obstacle into a manageable pedagogical approach (Bien and Mukherjee, 2024).

"Communication & Data Storytelling" not only ranked among the top skills but demonstrated increasing importance at higher job levels (Fig. 2), reflecting expanded leadership and management responsibilities. Similarly, "Machine Learning" and "Statistics and Math Knowledge" gained prominence with career progression, indicating their critical role in senior positions within the DSA landscape.

The top skills alone did not provide sufficient detail for comprehensive DSA skillset planning in tourism education; therefore, the granular skills were organized into **conceptual categories** (Fig. 3). *Managerial skills* category the most frequently required, often decisive for DSA success, aligns well with tourism education strengths. *Programming skills* category followed, with Python libraries dominating the Programming Libraries sub-area (Fig. 4).

Statistics and Machine Learning constitute major DSA education components. **Appendix B** details granular skill trends in this domain. Within machine learning, "Supervised Learning" demonstrated higher demand than "Deep Learning" and "Unsupervised Learning," while "Natural Language Processing" (NLP) emerged as a particularly sought-after specialization.

For Statistics (or maths), popular skills included "Experimentation," "A/B Testing," "Optimization," "Hypothesis Testing," "Regression," and "Bayesian Methods." However, educational planning should prioritize foundational concepts—Linear Algebra, Calculus, and Statistics—which provide comprehensive mathematical knowledge essential for DSA competency.

Big Data Management Skills (BDMS) are fundamental, with "Data Warehousing Knowledge," "Hadoop" or "Spark" ecosystems, and "Cloud Computing" emerging as prominent requirements. *Cloud Computing* expertise requires proficiency in at least one major platform, including AWS, Google Cloud, Azure, Databricks, or similar services.

Additional domains—Software Skills, Programming Libraries, Domain Knowledge, and Cloud Computing—offer further guidance for DSA course design in tourism (Fig. 4).

4.3. The Teams and Job Focus Areas

The analysis thus far addresses the "what" and "how" of DSA integration in tourism education, necessitating exploration of the "why" and "where" through examination of job responsibilities and functions within teams and broader Job Focus Areas (JFAs). This investigation aims to strategically align DSA components in tourism education with practical industry applications while assisting current practitioners and businesses in developing targeted strategies for career and competitive advantage.

Contextual analysis reveals that "Accommodation Providers" maintained more teams than "OTAs" or "Review Platforms." Career progression opportunities existed across approximately 20 teams extending to managerial levels. Geographic distribution showed the US, UK, and Netherlands as major markets hosting 15–25 teams, while Thailand, India, China, and Switzerland emerged as developing markets with 7–10 teams. These insights illuminate DSA employment potential within tourism.

To deepen understanding of DSA applications in tourism, the 31 initially identified teams (Y-axis, Fig. 5) were conceptualized into five theoretical JFA categories. Through theoretical sampling, comprehensive descriptive JFA profiles were developed, encompassing job quantities, primary teams, job level distributions, and prominent educational, skill, and experience requirements, supplemented by exemplary quotes illustrating core responsibilities. These profiles serve education planners and industry

professionals seeking to understand DSA's tourism applications and current business deployment strategies.

Figure 6 visualizes document-variable trends across Job Focus Areas (JFAs), illustrating how educational qualifications, job levels, and hiring organization types vary across domains. These distributions reinforce the earlier finding that accommodation providers and OTAs dominate recruitment, while data-intensive roles are concentrated in marketing and data science functions.

Marketing

Marketing (N = 173), was the predominant focus area where a large portion of teams aligned, for example, this JFA included teams such as Product, Supply, Sales, Pricing, Customer, Experience, Personalisation, Forecasting, Search, and others. The majority of these positions were at entry level (58%) followed by senior level (21%), manager level (15%) and senior manager level (5%). The most commonly required experience was between two to five years, with a PhD or master's being the most frequently required educational qualification, followed by a bachelor's degree. More than half of these jobs were offered by accommodation providers (53%), followed by online travel agencies (40%) and others. The top key skills included Python, SQL, R, communication and data storytelling, machine learning, cross-functionality, and statistics and math knowledge. A few key examples of responsibilities were as follows:

Utilise ML and other statistical techniques, to continuously generate insights related to the performance of models, marketing campaigns and advise the team on potential improvements ... Develop key metrics for our product and marketing campaigns.

Building machine learning models to categorize our supply ... Building probabilistic record linkage models ... Designing optimization processes to improve feedback collection for our users.

Define key sales performance metrics and build ROI models ... Develop sales experimentation and attribution frameworks ... sales conversion pipeline, customer segmentation, and customer success.

Data Science and Analytics

The second most common job focus area was Data Science and Analytics (N = 105), it mostly comprised teams such as Data Science, Data, Analytics, Machine Learning, Algorithms, Experimentation, Automation, and others. This area also followed similar job level trends i.e., most jobs at the entry level (47%), followed by senior level (34%), manager or director level (11%), and senior manager level (8%). In terms of qualifications, a PhD or Master's degree remained the most commonly required, followed by a master's degree, along with two to five years of experience. Online travel agencies (39%) and accommodation providers (29%) were the two main hiring organisations in this category, followed by transportation (15%), review platforms (11%), and miscellaneous (6%). The top skills required in this area included Python, communication and storytelling, SQL, R, machine learning, statistics and math, and the Hadoop ecosystem. Some examples of key responsibilities include the following:

[Work with] amazing amount of data as raw material for your projects, including over 780 million collected reviews and opinions, 5 terabytes of log data per day, and 300,000 photos and videos per day.

Translate data into metrics, dashboards, and analysis ... work collaboratively ... Identify critical business problems and create analytical/modelling framework and solutions ... Communicate analyses and recommendations to cross-functional stakeholders for decision making ... Training and mentoring other data scientists and analysts.

Financial Operations

Financial operations (N = 28), comprised teams that were focused on financial-cum-monetary aspects, it primarily included teams such as Payments, Finance, Insurance, Pricing, and Sales. More than half of these positions were offered at entry level (54%), followed by senior level (21%), manager or director level (21%), and senior manager levels (4%). A PhD or Master's was the most required degree followed by a Master's degree. The most frequently preferred experienced levels were between three to five years. About 71% of these jobs were offered by accommodation providers, followed by online travel agencies (21%) and review platforms (2%). The top key skills demanded in this area were communication and data storytelling, SQL, Python, R, statistics and math knowledge, data-driven problem solving, and cross-functionality. Some key examples of responsibilities included the following:

To identify and execute scalable ways to improve our platform's payment experience and to bring a scientific approach to decision making.

Focus on our wallet and rewards products ... Develop key metrics for the usage of our wallet and rewards products

Apply actuarial principles to evaluate and enhance insurance product offerings and pricing strategies ... Manage and optimize our claims fund, collaborating as needed with external actuaries from partner insurance organizations

Cyber Security

Cyber security (N = 22), mainly comprised the team "Trust / Risk and Security", including a few instances of teams such as Algorithms and Analytics. About 59% of these jobs were at entry level followed by senior level (18%), manager or director level (18%), and senior manager level (5%). In Cyber Security, a PhD or Master's was the only required degree in about 50% of the postings while the other postings did not mention any specific degree requirements. The experience level was relatively high in this area, ranging between two to eight years. In terms of hiring organisation type, the accommodation providers covered about 91% of the postings followed by online travel agencies (9%). The top key skills in this area were R, Python, communication and data storytelling, cross-functionality, SQL, data warehousing and ETL, and machine learning. A few key responsibilities of this area can be quoted as follows:

Building machine learning models to detect high risk activities like account takeovers, fake contents and fraudulent transactions, or high-risk entities like fake accounts or stolen cards ... Building NLP models to detect spam and inappropriate content on the fly

Collaborate with a strong team of engineers, product managers, designers and operation agents to build scalable and robust systems to detect and mitigate real-world safety risks

Define cutting edge experimentation and analytics to quantify the impact of trust product & operations

Human Resource and Management

The Human Resource and Management (N = 18) focus area included teams such as People, Strategy, Policy, and Product. More than half of these jobs were offered at the entry level (56%) followed by senior level (33%) and manager/director level (11%). The most frequently required degree was a PhD or Master's followed by a Master's Degree and Bachelor's degree or higher. The most commonly required experience ranged between two to six years. Accommodation providers (78%) and online travel agencies (22%) were the key recruiters in this area. The top key skills in this area were Python, SQL, R, communication and data storytelling, ownership mindset, data visualisation, and cross-functionality. A few key responsibilities of this area can be quoted as:

Work cross-functionally with Data Scientists and HR business partners to translate insights to action

Develop qualitative and quantitative methods for understanding talent acquisition and employee development, and measures of employee experience ... Evaluate and assess effectiveness of interview process, questions, and metrics ... Measure the impact of new methods for talent acquisitions and recruiting.

Using AI/ML models to derive insights from legislation and other unstructured data sources ... Collaboration with academics and experts on policy research topics.

These findings demonstrate that DSA is no longer confined to technical departments but is diffused across marketing, finance, security, and human-resource functions, signalling a structural transformation in tourism organizations. The next section discusses the implications of this shift for theory, management practice, and curriculum design.

5. Discussion

Prior research has increasingly emphasized the critical relevance of digital skills in the tourism industry, especially advanced competencies such as Data Science and Analytics (DSA). Once perceived as peripheral, these skills are now central to everyday operations, highlighting an urgent need to bridge the gap between industry and academia. Acknowledging this trend, leading educational institutions are already pioneering the integration of DSA into tourism curricula, a trajectory likely to be emulated

globally. Consequently, domain-specific understanding of DSA is indispensable for future educational strategies.

This study contributes significantly by empirically investigating DSA-related job roles in tourism, clarifying the specific skills and competencies demanded by the industry. Given the complexity and diversity of these skills, the conceptual framework developed herein provides a structured approach for integrating DSA effectively into tourism education. Additionally, businesses can leverage these insights to better align their strategic and operational needs with emerging DSA trends.

5.1 Theoretical Contributions

Research on digital skills in tourism frequently identifies gaps, emphasizing the need for more granular and empirically validated insights (Carlisle et al., 2023; Lamest & Brady, 2019; Minor et al., 2024). By focusing specifically on DSA, this study provides a detailed empirical understanding of critical competencies previously underexplored in broader digital skills literature. This empirically grounded approach directly addresses limitations noted in prior conceptual literature reviews (Adeyinka-Ojo et al., 2020; Ogbeide et al., 2021; Sharma, 2024) and significantly advances discussions around aligning educational practices with real-world industry demands.

5.2 Methodological Contributions

Methodologically, this study employed Job Advertisement Analysis (JAA), an approach proven effective for advancing tourism scholarship (Arcodia and Barker, 2002; Bernstein and Woosnam, 2019; Robinson et al., 2010), though Arcodia, Novais, and Le (2020) noted its underutilization in tourism research.

This study enhanced the JAA approach by incorporating Grounded Theory, enabling findings to authentically reflect real-world industrial trends rather than conforming to predetermined theoretical frameworks. For future applications, this research recommends emphasizing methodological rigor in contextual positioning, memo-writing, coding processes, and theoretical sampling conditions, as detailed in the methodology section.

This JAA-Grounded Theory integration addresses the documented lack of industrial input in tourism education planning (Airey and Nightingale, 1981; Hsu et al., 2017; Tribe, 2002), offering a robust framework for bridging academic theory with industry practice.

5.3 Practical Contributions

This study illuminates the nature of DSA in tourism through job, skill, and educational perspectives, establishing "what" requires incorporation into tourism education at both granular and conceptual levels. However, understanding "how" to integrate these skills and concepts into tourism education remains paramount. Because multiple skills and techniques often achieve similar outcomes with overlapping learning components, strategic arrangement becomes essential for effective pedagogy.

Technical Integration Approach

- Technical skills and theoretical concepts should be structured to promote symbiotic understanding. Mathematical principles can be implemented through programming languages on big data platforms using cloud computing and databases under the umbrella of machine learning. This approach enables simultaneous learning across mathematical foundations, programming techniques, software applications, cloud computing, and machine learning methodologies.
- Programming skills may initially challenge social science and business-focused learners (Bien and Mukherjee, 2024), yet Large Language Models (LLMs) are transforming this landscape. Programming is becoming as accessible as writing "English prompts," though basic syntax understanding remains crucial. While digital skills literature considers programming less important (Carlisle et al., 2023), from a data science perspective, it remains vital for developing complex innovative solutions unavailable in standard software packages.
- Managerial skills demonstrate strong alignment with previous research, including communication (Carlisle et al., 2023; Minor et al., 2024), dashboard creation, data-driven decision-making, and cross-functional capabilities (Lamest and Brady, 2019), plus critical thinking, problem-solving, interpersonal relationships, and teamwork (Adeyinka-Ojo et al., 2020). These skills require minimal additional integration due to their established presence in existing tourism education frameworks.

Curricular Integration Strategy

The question of whether to introduce separate subjects or integrate modules within existing syllabi can be addressed through this study's Job Focus Areas (JFAs) framework. JFAs provide nuanced understanding of data science applications in tourism, with many JFA designations corresponding to existing tourism course modules and business departments. This alignment suggests that DSA integration can occur within current curricula rather than requiring entirely new modules (Martin-Rios et al., 2017; Ogbeide et al., 2021).

Tourism marketing courses can incorporate electronic word-of-mouth (e-WOM) analysis using web scraping, sentiment analysis, and topic modelling techniques. This integration aligns with Carlisle et al.'s (2023) identification of "monitoring online reviews" as essential digital competency and Lamest and Brady's (2019) emphasis on customer data interpretation skills for tourism marketing managers. Financial Management modules can introduce technical concepts supporting automated revenue systems understanding (Haynes and Egan, 2017). Human Resources and Management courses can integrate HR analytics and people analytics concepts (Martin-Rios et al., 2017).

Cybersecurity represents a genuinely new curricular requirement. Given its growing importance, tourism professionals cannot afford theoretical ignorance in this domain. Werthner (2022) emphasizes that "cybersecurity and data protection will become permanent topics of our society." Carlisle et al. (2023) identify increasing importance of cybersecurity, data protection, and ethics skills in tourism. Concerningly, this research found limited demand for "data ethics" skills, highlighting a critical gap for DSA's future in tourism.

6. Conclusion

The analysis of 346 tourism-industry job advertisements exposes a decisive, sector-wide shift toward data-science and analytics (DSA) capability. Accommodation providers and online travel agencies (OTAs) alone generated 86% of the vacancies, signalling that the largest value-chain actors are already reorganising around data-driven decision making. More than half of all roles (54%) were entry-level, confirming that demand begins at the graduate tier and scales upward through senior specialist and leadership positions. The discussion interprets these patterns through four lenses—theory, managerial practice, curriculum design and policy/future research—to show how the findings advance knowledge and inform action.

This study is subject to several limitations inherent in its methodological design. Primarily, the exclusive analysis of English-language job advertisements means the findings are most representative of anglophone markets, such as the US and UK, and likely understate the demand and specific skill configurations in major non-English-speaking tourism economies. Second, job advertisements articulate an employer's ideal candidate profile, which may not perfectly mirror the competencies of the individual ultimately hired; this method also omits positions filled through internal promotions or confidential headhunting. Finally, the five-month data collection window offers a valuable but static snapshot of a rapidly evolving field. A longitudinal analysis would be necessary to track the dynamic evolution of DSA skill requirements over time. Despite these constraints, the study provides a robust and empirically grounded foundation for future research.

6.1 Theoretical implications

By locating granular skill signals inside an inductively generated Job Focus Area (JFA) schema, these findings imply that tourism human-capital theory must evolve beyond its traditional service-orientation. Results demonstrate that advanced technical capital (e.g., Python, SQL, cloud data-warehousing) and higher-order cognitive capital (communication, data storytelling) are now complementary, not substitute, assets in tourism firms. This complements and refines the “skills-gap” discourse: rather than a binary digital/non-digital divide, the gap is layered—technical proficiency is a necessary but insufficient condition for progression unless paired with narrative and leadership competences. Conceptually, the GT-enhanced job-advertisement method also operationalises labour-market signalling theory for tourism research, showing how employers broadcast evolving competence hierarchies in real time.

6.2 Managerial implications

Talent strategy

HR leaders in accommodation and OTA firms should recalibrate job architecture and progression criteria. Entry-level tracks can justifiably weight technical benchmarks (e.g., SQL fluency, experimentation pipelines), but promotion gateways must include demonstrable communication and cross-functional influence, echoing the observed rise of “data storyteller” language at senior bands.

Upskilling and retention: Given the thin external supply of senior-level tourism data scientists, firms should institute internal academies that ladder staff from analyst to specialist to leader. The JFA map supplies a ready-made scaffold: Marketing Analytics, Revenue & Pricing, Cyber-Security Analytics, Financial Operations Analytics and HR/People Analytics each map onto recognisable business units, enabling targeted learning journeys rather than generic “big-data” workshops.

Technology investment

The dominance of cloud-native tooling in adverts implies that capital expenditure on on-premises BI stacks will accelerate obsolescence. Vendor selection should privilege open-source-friendly ecosystems (e.g., Databricks, Snowflake) that align with the Python/R skill profile sought in the labour market.

6.3 Educational implications (curriculum design)

Module integration vs. standalone courses. The JFA evidence shows clear affinities between DSA tasks and existing tourism subjects (e-WOM analysis in Marketing; dynamic pricing in Revenue Management; HR analytics in People Management). Therefore, most institutions can embed DSA within current modules rather than building entirely new degree pathways—an approach that mitigates contact-hour inflation and staff-load constraints.

Progressive skill scaffolding. A three-stage roadmap is recommended:

1. Foundations (Year 1–2). Statistics, linear algebra, SQL, data ethics.
2. Applications (Year 2–3). Domain-anchored labs (marketing sentiment analysis, demand-forecasting simulators).
3. Integration (Capstone). Cross-functional projects using real corporate datasets under NDAs, evaluated jointly by academic and industry mentors.

Assessment innovation. Replace closed-book exams with portfolio-based assessment: annotated notebooks, interactive dashboards and reflection essays on data-driven decision impact. This aligns with employer expectations for demonstrable, job-ready artefacts.

6.4 Policy and future-research implications

National tourism bodies and accreditation agencies should recognise DSA capability as a quality-assurance metric. Scholarships and industry secondments that reward interdisciplinary DSA study will accelerate pipeline development, particularly in non-English-speaking markets that were beyond this study’s language scope.

Future research can:

- Replicate the JAA-GT approach in Mandarin, Spanish or Arabic job boards to surface culturally specific DSA profiles.

- Conduct longitudinal JAA sweeps to track how emergent technologies—generative AI agents, privacy-preserving analytics—reconfigure the skill hierarchy.
- Experiment with curriculum-action research: embed the proposed roadmap in pilot programmes and measure graduate outcomes relative to control cohorts.

In sum, this study demonstrates that tourism's data turn is both pervasive and multi-dimensional, requiring coherent responses from theorists, enterprise leaders, educators and policymakers alike. By translating real-time labour-market signals into actionable frameworks, the study offers a pragmatic scaffold for synchronising academic preparation with industry evolution—closing, at last, the long-lamented academy–industry divide in tourism education.

Declarations

Author Contribution

Both authors contributed equally.

References

1. Adeyinka-Ojo S, Lee S, Abdullah SK et al (2020) Hospitality and tourism education in an emerging digital economy. *Worldw Hospitality Tourism Themes* 12(2):113–125
2. Airey D, Nightingale M (1981) Tourism occupations, career profiles and knowledge. *Annals Tourism Res* 8(1):52–68
3. Arcodia C, Barker T (2002) A review of web-based job advertisements for Australian event management positions. *J Sport Tourism* 7(3):23–25
4. Arcodia C, Novais MA, Le TH (2020) Using Job Advertisements to Advance Event Management Research. *Event Manage* 24(5):661–664
5. Barnes SJ, Kirshner SN (2021) Understanding the impact of host facial characteristics on Airbnb pricing: Integrating facial image analytics into tourism research. *Tour Manag* 83:104235
6. Bernstein JD, Woosnam KM (2019) Same same but different: Distinguishing what it means to teach English as a foreign language within the context of volunteer tourism. *Tour Manag* 72:427–436
7. Bien J, Mukherjee G (2024) Generative AI for Data Science 101: Coding Without Learning To Code. *J Stat Data Sci Educ* : 1–12
8. Bilotta E, Bertacchini F, Gabriele L et al (2021) Industry 4.0 technologies in tourism education: Nurturing students to think with technology. *J Hospitality Leisure Sport Tourism Educ* 29:100275
9. Buhalis D, Law R (2008) Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tour Manag* 29(4):609–623
10. Buhalis D, Harwood T, Bogicevic V et al (2019) Technological disruptions in services: lessons from tourism and hospitality. *J Service Manage* 30(4):484–506

11. Busulwa R, Pickering M, Mao I (2022) Digital transformation and hospitality management competencies: Toward an integrative framework. *Int J Hospitality Manage* 102:103132
12. Busulwa R, Pickering M, Pathiranage NW (2024) Readiness of hospitality and tourism curricula for digital transformation. *J Hospitality Leisure Sport Tourism Educ* 35:100519
13. Carlisle S, Ivanov S, Dijkmans C (2023) The digital skills divide: evidence from the European tourism industry. *J Tourism Futures* 9(2):240–266
14. Charmaz K (2006) *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Contructing Grounded Theory. SAGE, London
15. Choi Y, Rasmussen E (2009) What Qualifications and Skills are Important for Digital Librarian Positions in Academic Libraries? A Job Advertisement Analysis. *J Acad Librariansh* 35(5):457–467
16. Doborjeh Z, Hemmington N, Doborjeh M et al (2022) Artificial intelligence: a systematic review of methods and applications in hospitality and tourism. *International Journal of Contemporary Hospitality Management* 34(3). Emerald Group Holdings Ltd.: 1154–1176
17. Elnasr Sobaih A, Jones E (2015) Bridging the hospitality and tourism university–industry research gap in developing countries: The case of Egypt. *Tourism Hospitality Res* 15(3):161–177
18. Fuchs M, Höpken W, Lexhagen M (2014) Big data analytics for knowledge generation in tourism destinations – A case from Sweden. *J Destination Mark Manage* 3(4):198–209
19. Glaser BG (1992) *Emergence Vs Forcing: Basics of Grounded Theory Analysis*. Emergence vs. forcing. Mill Valley, CA: Sociology Press
20. Glaser BG, Strauss AI (1967) *The Discovery of Grounded Theory. Strategies for Qualitative Research*. Aldine, Chicago
21. Haynes N, Egan D (2017) Revisiting the relevance of economic theory to hotel revenue management education and practice in the era of Big Data. *Res Hospitality Manage* 7(1):65–73
22. Hsu CHC (2018) Tourism education on and beyond the horizon. *Tourism Manage Perspect* 25:181–183
23. Hsu CHC, Xiao H, Chen N (2017) Hospitality and tourism education research from 2005 to 2014. *Int J Contemp Hospitality Manage* 29(1):141–160
24. Ivanov S (2020) The impact of automation on tourism and hospitality jobs. *Inform Technol Tourism* 22(2):205–215
25. Lamest M, Brady M (2019) Data-focused managerial challenges within the hotel sector. *Tourism Rev* 74(1):104–115
26. Li J, Xu L, Tang L et al (2018) Big data in tourism research: A literature review. *Tourism Management* 68. Elsevier Ltd: 301–323
27. Lv H, Shi S, Gursoy D (2022) A look back and a leap forward: a review and synthesis of big data and artificial intelligence literature in hospitality and tourism. *Journal of Hospitality Marketing and Management* 31(2). Routledge: 145–175

28. Mariani M (2020) Big Data and analytics in tourism and hospitality: a perspective article. *Tourism Rev* 75(1):299–303
29. Mariani M, Baggio R, Fuchs M et al (2018) Business intelligence and big data in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management* 30(12). Emerald Publishing Limited: 3514–3554
30. Martin-Rios C, Pougnet S, Nogareda AM (2017) Teaching HRM in contemporary hospitality management: a case study drawing on HR analytics and big data analysis. *J Teach Travel Tourism* 17(1):34–54
31. Maxwell JA (2013) *Qualitative Research Design: An Interactive Approach*. SAGE
32. McCartney G, Kwok SSH (2023) Closing the Gap between Hospitality Industry and Academia Research Agendas: An Exploratory Analysis Case Study. *J Qual Assur Hospitality Tourism* 24(2):192–210
33. Meyer MA (2019) Healthcare data scientist qualifications, skills, and job focus: a content analysis of job postings. *J Am Med Inform Assoc* 26(5):383–391
34. Minor K, McLoughlin E, Carlisle S (2024) The Digital Skills Gap – Is it Time to Rethink the Needs of Tourism and Hospitality Organizations in the UK? *J Hospitality Tourism Educ* : 1–12
35. Mohanty S, Mohanty A (2019) A skill- gap study: An analytical approach with a special focus on Tourism Education and the Tourism Industry in Odisha. *Afr J Hospitality Tourism Leisure* 8(3)
36. Ogbeide G-C, Fu Y-Y, Cecil AK (2021) Are hospitality/tourism curricula ready for big data? *J Hospitality Tourism Technol* 12(1):112–123
37. Patten ML, New (2017) *Understanding Research Methods -an Overview of the Essentials*, 10th edn. Routledge, New York
38. Ralph N, Birks M, Chapman Y (2014) Contextual Positioning. *SAGE Open* 4(3):215824401455242
39. Robinson RNS, Arcodia CV, Tian C et al (2010) Cookery-related employment vacancy advertisements and skills shortages. *Int J Contemp Hospitality Manage* 22(4):483–499
40. Sharma S (2023) Rise of Big Data in the Industry 4.0 Era and the Crucial Need for Data Science in Sustainable Tourism. In: *Impact of Industry 4.0 on Sustainable Tourism*. Emerald Publishing Limited, pp. 133–149. Available at: <https://www.emerald.com/insight/content/doi/10.1108/978-1-80455-157-820231009/full/html>
41. Sharma S (2024) Data Scientist in Tourism: Needs, Skills, and Scope. In: Sharma A (ed.) *International Handbook of Skill, Education, Learning, and Research Development in Tourism and Hospitality*. Singapore: Springer Nature Singapore, pp. 1–16. Available at: https://doi.org/10.1007/978-981-99-3895-7_16-1
42. Sigala M (2018) New technologies in tourism: From multi-disciplinary to anti-disciplinary advances and trajectories. *Tourism Manage Perspect* 25:151–155
43. Tribe J (2002) Research trends and imperatives in tourism education. *Acta Touristica* 14(1):61–81

44. Werthner H (2022) e-Tourism: An Informatics Perspective. In: Xiang Z, Fuchs M, Gretzel U, (eds) *Handbook of E-Tourism*. Cham: Springer. Available at: https://doi.org/10.1007/978-3-030-48652-5_1
45. Wickham H (2016) Ggplot2: Elegant Graphics for Data Analysis. Springer-, New York
46. Wickham H, Averick M, Bryan J et al (2019) Welcome to the Tidyverse. *J Open Source Softw* 4(43):1686
47. Xiang Z, Schwartz Z, Gerdes JH et al (2015) What can big data and text analytics tell us about hotel guest experience and satisfaction? *Int J Hospitality Manage* 44:120–130
48. Xiang Z, Du Q, Ma Y et al (2017) A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tour Manag* 58:51–65
49. Zhan L, Cheng M, Zhu J (2024) Progress on image analytics: Implications for tourism and hospitality research. *Tour Manag* 100:104798
50. Zhu J, Cheng M, Wang Y, Wendy) (2024) (Viewer In-Consumption Engagement in Pro-Environmental Tourism Videos: A Video Analytics Approach. *Journal of Travel Research*. Epub ahead of print 10 January 2024. 10.1177/00472875231219634

Figures

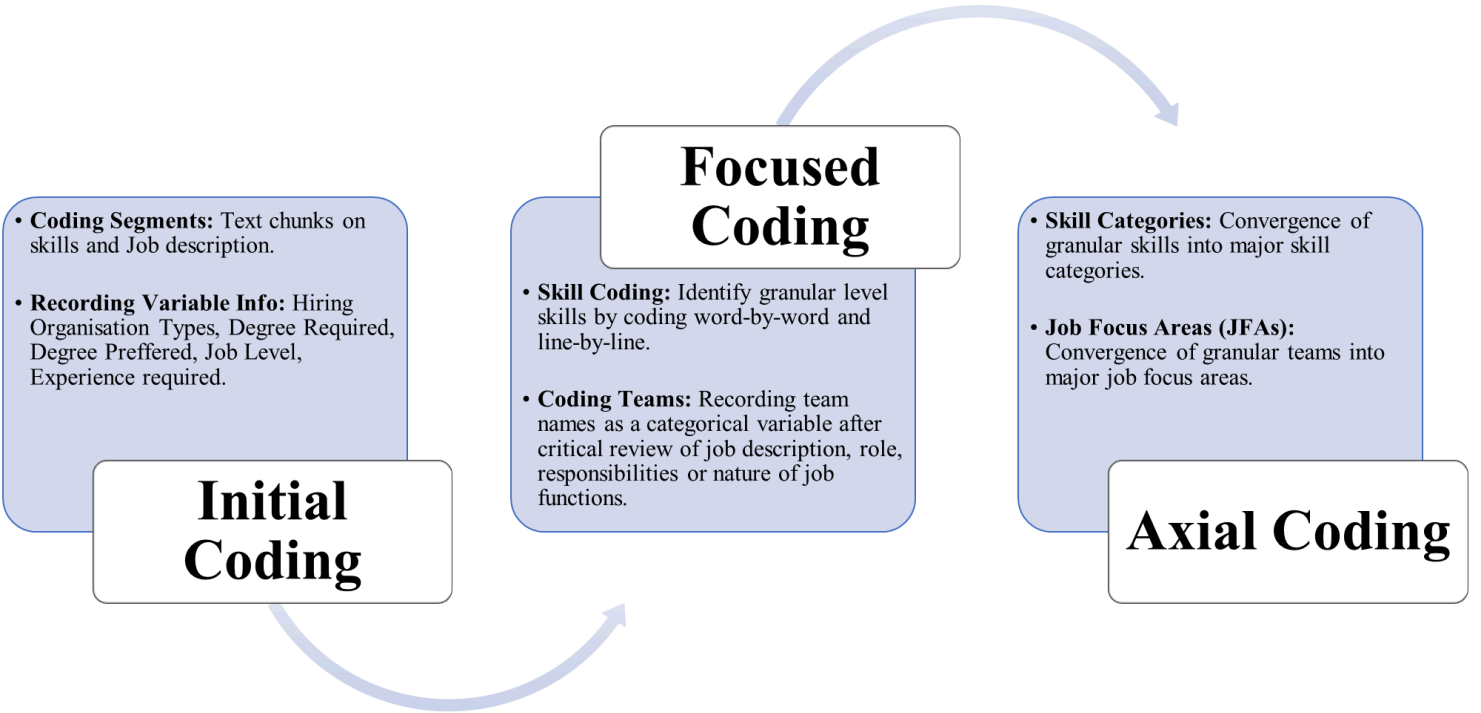


Figure 1

Qualitative Coding Workflow

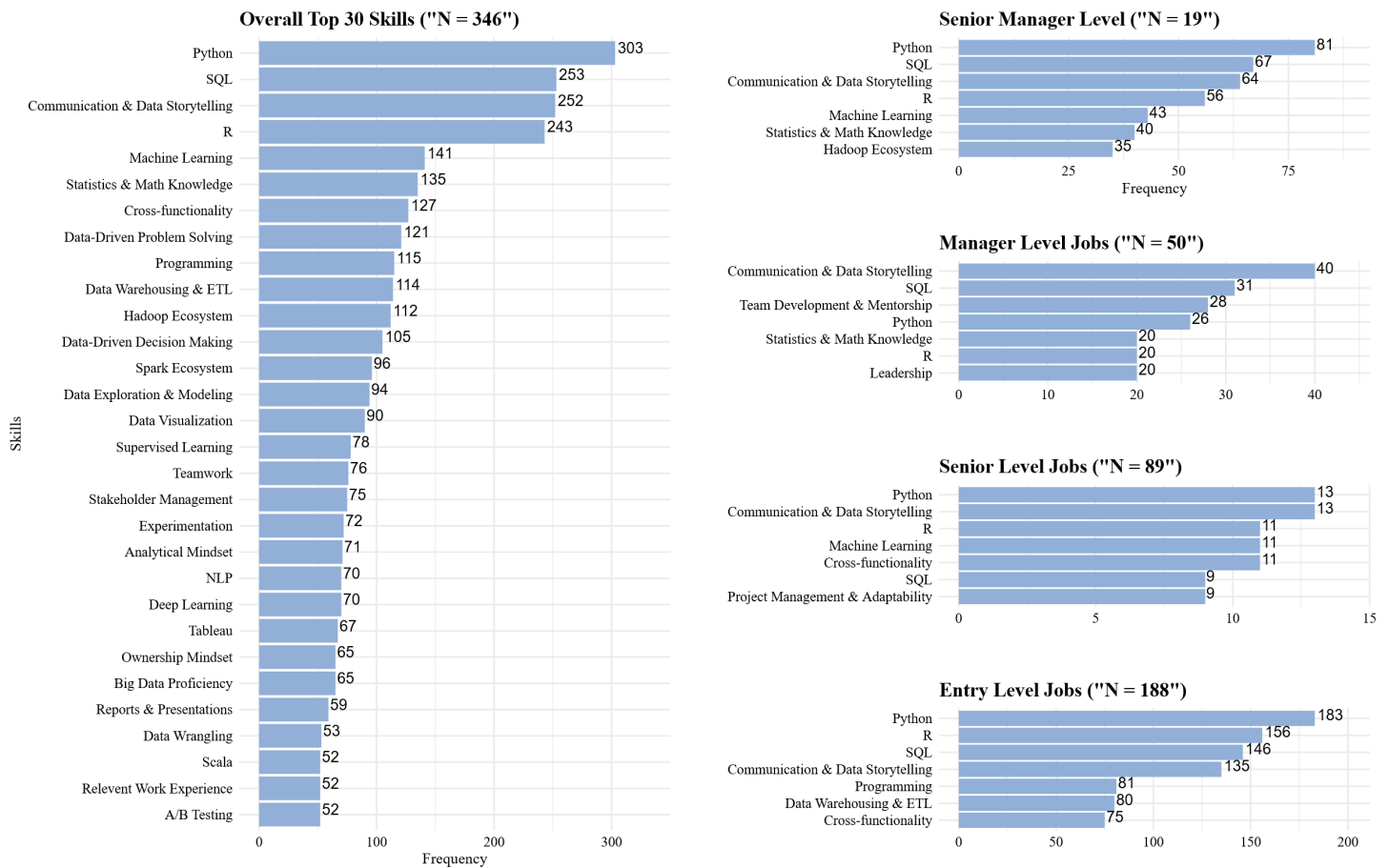


Figure 2

Distribution of the top 30 skills in the overall dataset and the top seven skills across job levels.

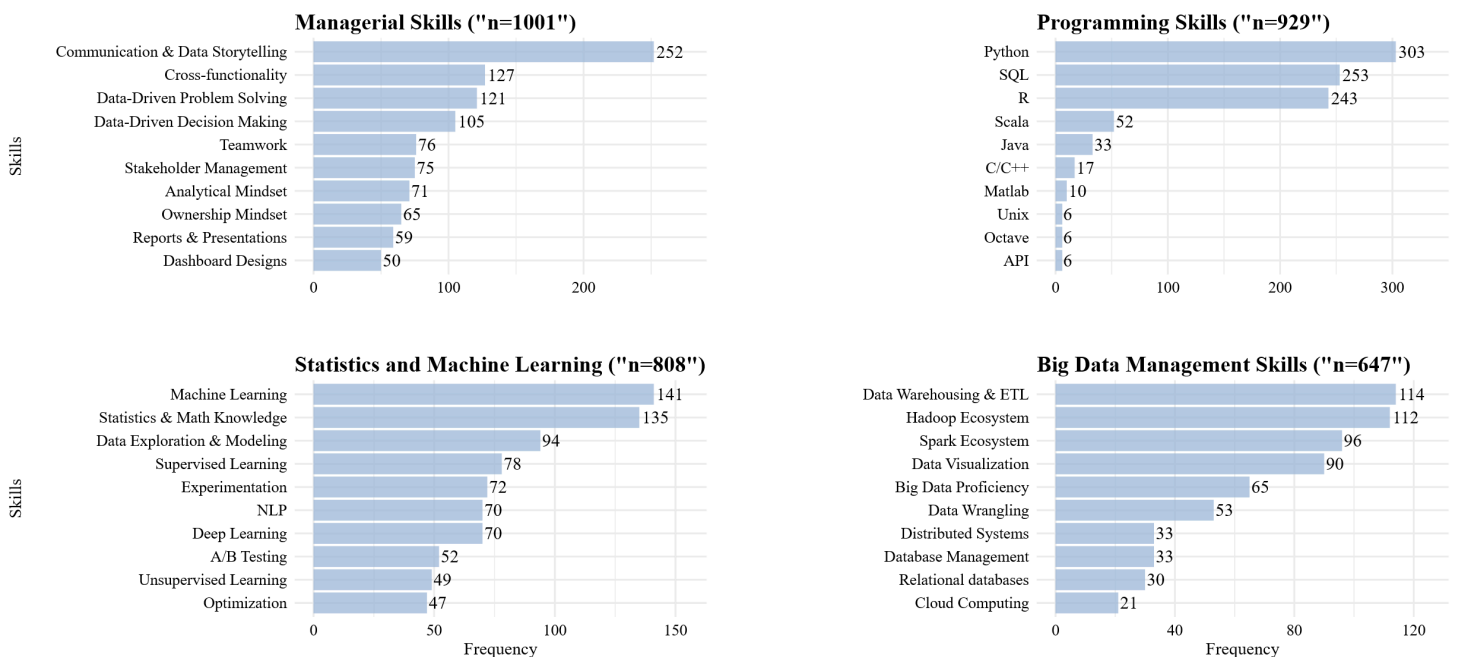


Figure 3

Major Skill Categories

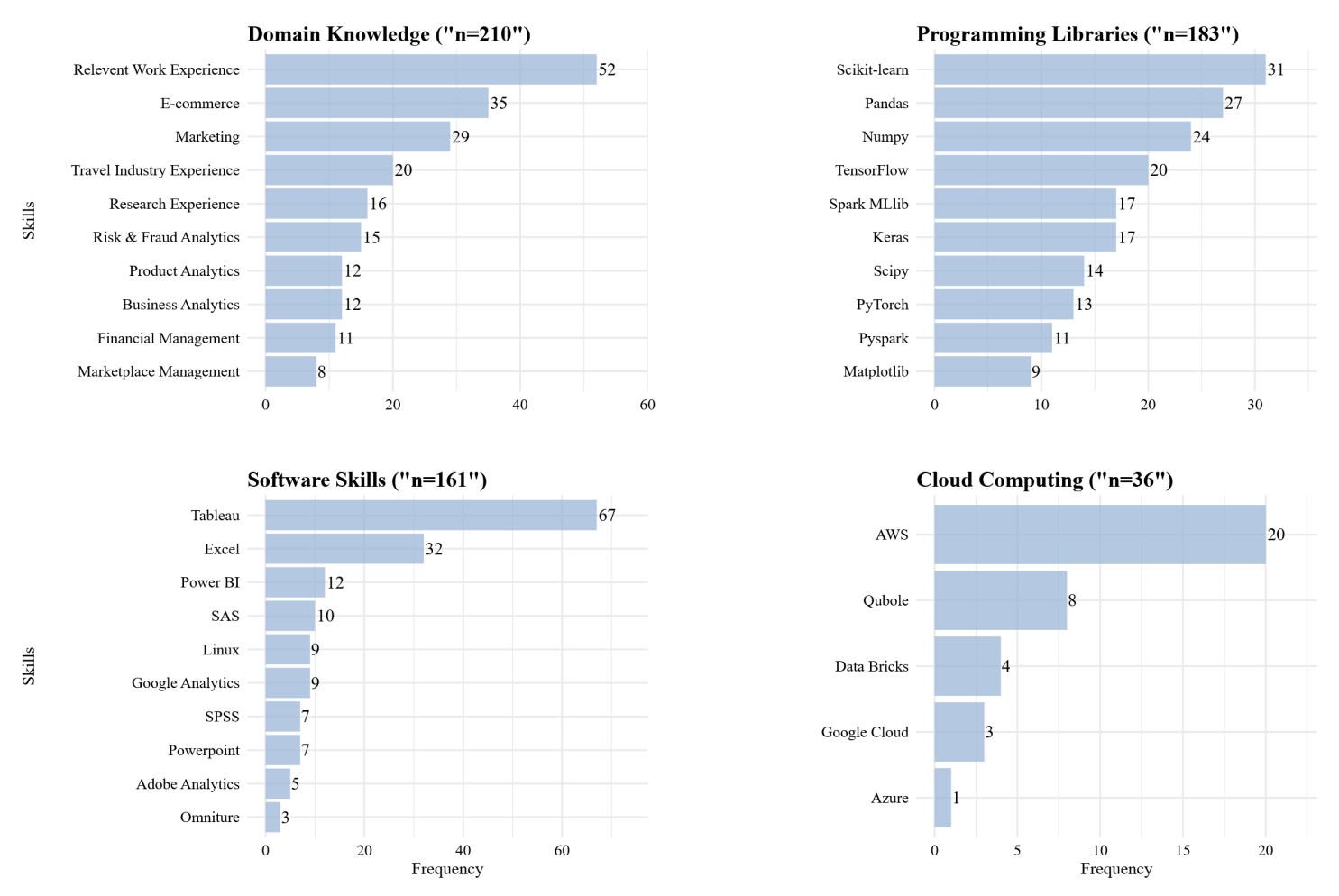


Figure 4

Miscellaneous Skill Categories.

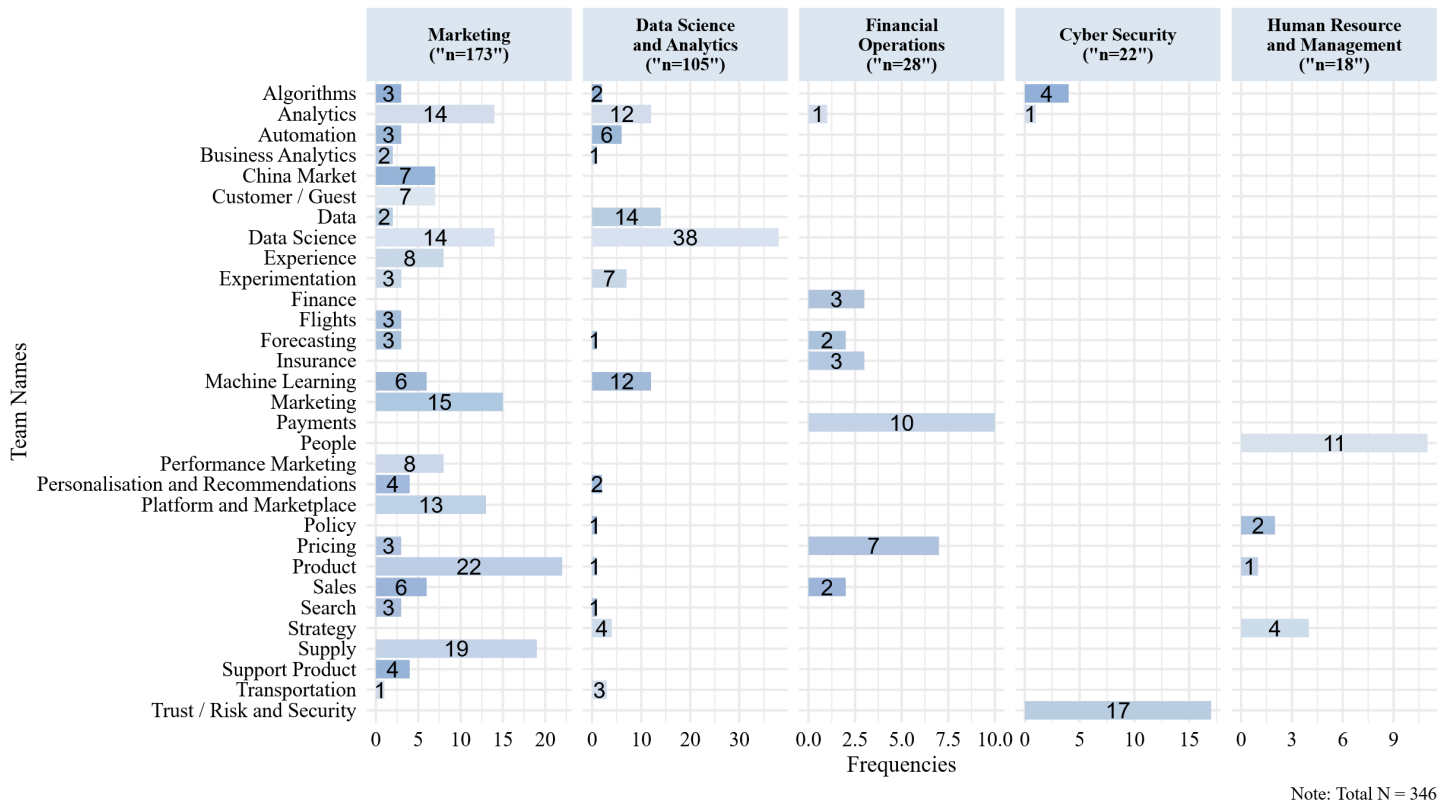


Figure 5

Distribution of teams across Job Focus Areas.

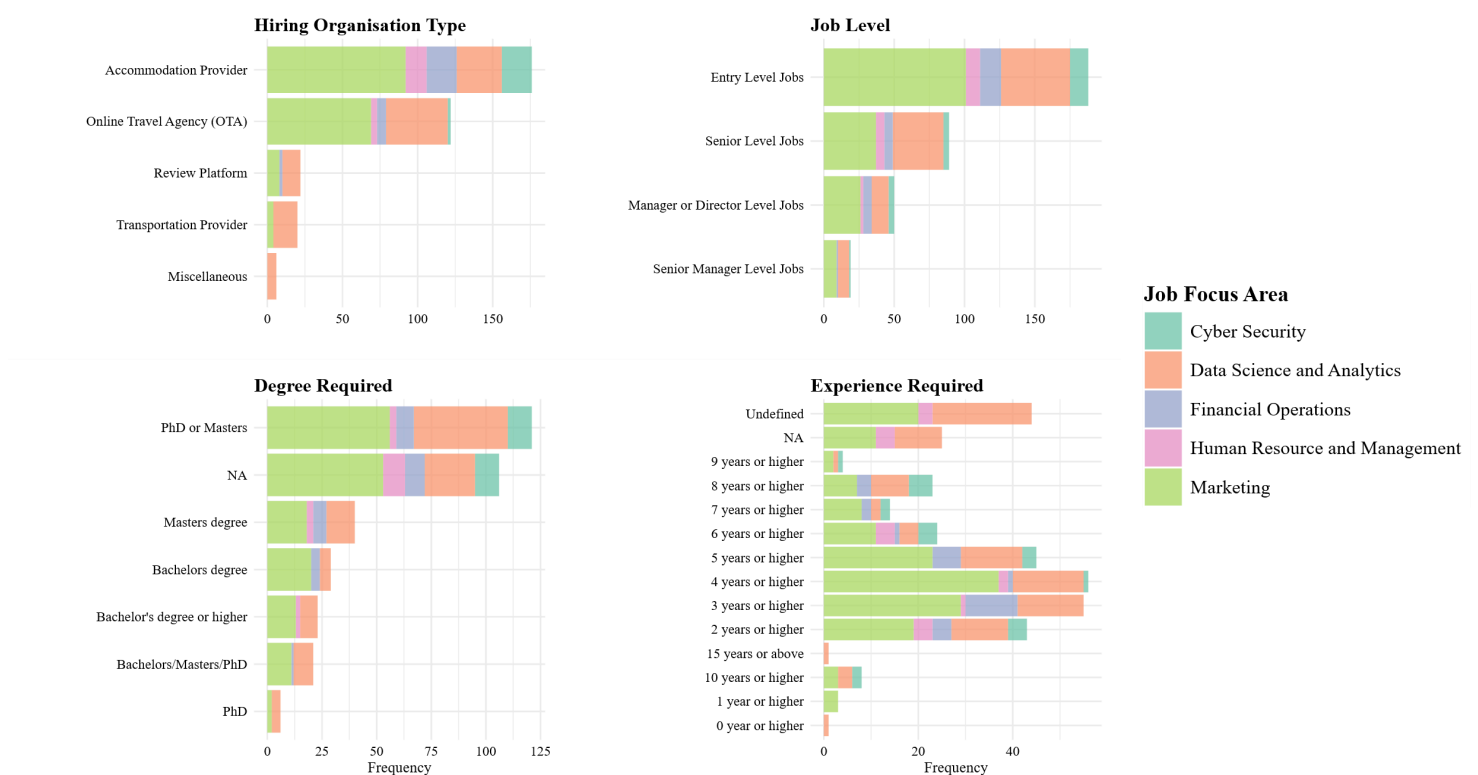


Figure 6

Document variable trends across job focus areas.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AppendixA.docx](#)
- [AppendixB.svg](#)