Annex A: OfS analysis of Designated Data body (DDB) data for Artificial Intelligence and Data Science postgraduate conversion course scholarship programme 2023-24

Postgraduate conversion course programme: Characteristics analysis

This annex provides an analysis of the artificial intelligence (AI) and data science postgraduate conversion course (PGCC) scholarship programme conducted by the Office for Students (OfS). The purpose of the analysis is to assess the impact of the programme on increasing enrolment of students from disadvantaged backgrounds onto postgraduate courses in data science and AI. The broader goal is to evaluate whether the programme achieved its aim to increase the diversity of the AI and data science workforce.

The designated data body student record has been used to examine the characteristics of entrants to PGCC courses in the 2023-24 academic year. This is to aid the evaluation of the programme by CRAC. Characteristic information has been collected for four groups of students:

- entrants on a PGCC course in receipt of a programme scholarship
- entrants on a PGCC course not in receipt of a programme scholarship
- all entrants on a PGCC course
- entrants on a comparison group of courses identified as being in the fields of AI or data science at master's level that do not have any scholarship students.

PGCC courses have been identified by using course initiative ID (courseinitid) taking value '031' in the Course Initiatives dataset, and students in receipt of a programme scholarship have been identified through student initiative ID (stuinitid) taking value '034' in the Student Initiatives dataset. The quality of this data for some participating providers was low and further primary data collection was conducted to improve consistency and reliability. The final dataset contained courses from 28 higher education providers.

A sampling frame of non-PGCC courses in the fields of AI and data science has been identified in the student data by searching for relevant keywords in course titles, restricted to masters' level programmes. To construct the comparison group, a stratified non-random sample of courses has been selected to ensure that the distribution of students across high, medium, and low-tariff providers (as defined by the OfS classification[1]) approximately matched that of the PGCC courses with at least one scholarship student.

[1] See OfS, 'Provider typologies 2022: Methodology for grouping OfS-registered providers'.

Characteristics of entrants to artificial intelligence and data science courses

The characteristic proportions reported here are for all students:

- age (banded)
- disability
- mode of study
- domicile (UK or non-UK)
- sex.

These characteristic proportions reported are limited to UK-domiciled students:

- ethnicity
- first degree subject
- science, technology, engineering and mathematics (STEM) categorisation
- National Statistics Socioeconomic Classification (NS-SEC).

Linking of administrative data was used to establish information on first degree subject and NS-SEC by finding the most recent undergraduate record for those students educated in the UK. As this can still result in multiple records for a student, to resolve to one historical record only AI and data science subject areas have been taken as a preference, followed by cognate, STEM, and non-STEM subjects in that order.

The NS-SEC field is restricted to full-time undergraduate students applying through UCAS.

The summaries are contained in Tables A1 to A8 below. All numbers reported in the following tables are rounded to the nearest five, while all proportions are rounded to the nearest whole number. Due to this subtotals and totals may not sum exactly.

We estimate that there are around 12,550 students on data science and AI master's level courses. This estimate was arrived at by filtering to all masters' level programmes, both PGCC and non-PGCC, with a primary focus on either data science or AI.

Table A1: Entrants by age (banded)

Age	PGCC Scholarship students		PGCC Non-scholarship students		Total PGCC students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
25 or under	305	45%	1,795	50%	2,100	49%	3,105	65%
Over 25	380	55%	1,765	50%	2,145	51%	1,655	35%
Total	**685**	**100%**	**3,560**	**100%**	**4,245**	**100%**	**4,760**	**100%**

Table A2: Entrants by disability

Disability	PGCC Scholarship students		PGCC Non-scholarship students		Total students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
Not disabled	470	69%	3,350	94%	3,820	90%	4,490	94%
Disabled	215	31%	215	6%	430	10%	270	6%
Total	**685**	**100%**	**3,565**	**100%**	**4,245**	**100%**	**4,760**	**100%**

Table A3: Entrants by mode of study

Mode of study	PGCC Scholarship students		PGCC Non-scholarship students		Total PGCC students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
Apprenticeship	-	-	-	-	-	-	50	1%
Part-time	165	24%	535	15%	700	16%	380	8%
Full-time	520	76%	3,025	85%	3,545	84%	4,325	91%
Total	**685**	**100%**	**3,560**	**100%**	**4,245**	**100%**	**4,755**	**100%**

Table A4: Entrants by domicile (UK vs non-UK)

Domicile	PGCC Scholarship students		PGCC Non-scholarship students		Total PGCC students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
Not UK-domiciled	60	9%	2,745	77%	2,805	66%	3,815	80%
UK-domiciled	625	91%	815	23%	1,445	34%	940	20%
Total	**685**	**100%**	**3,560**	**100%**	**4,245**	**100%**	**4,755**	**100%**

Table A5: Entrants by sex

Sex	PGCC Scholarship students		PGCC Non-scholarship students		Total PGCC students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
Unknown	0	0%	0	0%	0	0%	5	0%
Male	240	34%	2,255	63%	2,495	59%	3,165	66%
Female	445	66%	1,300	36%	1,745	41%	1,590	33%
Other sex	0	0%	5	0%	5	0%	0	0%
Total	**685**	**100%**	**3,560**	**100%**	**4,245**	**100%**	**4,760**	**100%**

Table A6: Entrants by ethnicity (known ethnicity only)

Ethnicity	PGCC Scholarship students		PGCC Non-scholarship students		Total PGCC students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
Asian	115	19%	170	22%	285	20%	195	22%
Black	225	37%	120	15%	345	25%	120	14%
Mixed	50	8%	45	6%	95	7%	40	5%
Other	25	4%	30	4%	55	4%	45	5%
White	195	32%	425	54%	620	44%	475	54%
Total	**610**	**100%**	**790**	**100%**	**1,400**	**100%**	**875**	**100%**

Note: This data is restricted to UK-domiciled students only.

Table A7: Entrants by first-degree subject (known subject only)

Subject	PGCC Scholarship students		PGCC Non-scholarship students		Total PGCC students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
Agriculture, food and related studies	0	0%	0	0%	0	0%	0	0%
Architecture, building and planning	0	0%	5	1%	5	0%	0	0%
Biological and sport sciences	65	10%	70	9%	135	9%	30	4%
Business and management	115	18%	120	15%	235	16%	55	7%
Combined and general studies	0	0%	5	1%	10	1%	5	1%
Computing	40	6%	110	14%	155	11%	225	29%
Design, and creative and performing arts	15	2%	25	3%	40	3%	5	1%
Education and teaching	10	2%	15	2%	25	2%	5	1%
Engineering and technology	35	5%	70	9%	105	7%	65	8%
Geography, earth and environmental studies	15	2%	10	1%	25	2%	10	1%
Historical, philosophical and religious studies	15	2%	30	4%	40	3%	25	3%
Language and area studies	30	5%	40	5%	70	5%	10	1%
Law	25	4%	15	2%	45	3%	10	1%

Subject	PGCC Scholarship students		PGCC Non-scholarship students		Total PGCC students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
Mathematical sciences	35	5%	45	6%	80	6%	125	16%
Media, journalism and communications	5	1%	15	2%	20	1%	5	1%
Medicine and dentistry	10	2%	10	1%	20	1%	5	1%
Physical sciences	30	5%	55	7%	80	6%	55	7%
Psychology	30	5%	20	3%	50	3%	15	2%
Social sciences	100	15%	65	8%	165	12%	80	10%
Subjects allied to medicine	75	12%	60	8%	130	9%	45	6%
Veterinary sciences	0	0%	0	0%	0	0%	0	0%
Total	**650**	**100%**	**785**	**100%**	**1,430**	**100%**	**775**	**100%**

Note: This data is restricted to UK-domiciled students only.

Table A8: Entrants by STEM categorisation (known subject only)

STEM category	PGCC Scholarship students		PGCC Non-scholarship students		Total PGCC students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
Cognate	45	6%	145	14%	190	11%	225	31%
far STEM	200	30%	200	22%	400	26%	95	13%
core STEM	100	15%	225	22%	325	19%	240	33%
non-STEM	335	49%	375	42%	710	45%	175	24%
Total	**650**	**100%**	**780**	**100%**	**1,440**	**100%**	**735**	**100%**

Note: This data is restricted to UK-domiciled students only.

Table A9: Entrants by NS-SEC (known socioeconomic status only)

NS-SEC	PGCC Scholarship students		PGCC Non-scholarship students		Total PGCC students		Comparison group AI and data science students	
	Number	%	Number	%	Number	%	Number	%
Higher managerial & professional occupations	55	16%	125	29%	180	23%	125	31%
Lower managerial & professional occupations	85	24%	115	27%	200	26%	95	23%
Intermediate occupation	65	19%	50	12%	115	15%	40	10%
Small employers & own account workers	30	9%	25	6%	55	7%	30	7%
Lower supervisory & technical occupations	25	7%	15	4%	35	5%	25	6%
Semi-routine occupations	55	16%	55	13%	110	14%	55	14%
Routine occupations	35	10%	40	9%	75	10%	35	9%
Never worked & long-term unemployed	0	0%	0	0%	0	0%	0	0%
Total	**350**	**100%**	**425**	**100%**	**770**	**100%**	**405**	**100%**

Note: This data is restricted to UK-domiciled students only.