Social care workforce report

2022

Index

1. Introduction

This report sets out the key findings of our second annual workforce data collection, which we carried out in 2022.

After talking to our partners and taking on board what we learned from the first year of collection, we made some changes to the collection process in 2022. We decided to collect data directly from both local authorities and care providers, rather than ask local authorities to collect data from providers on our behalf. Local authorities are now asked to provide their internal data and a list of providers for reference.

This year, we contacted all independent providers of social care in Wales to ask for data about their workforce. We hoped these changes would reduce the burden on local authorities and allow us to have a single, co-ordinated approach that would improve data quality.

These changes to the process meant we needed to delay the start of collection to allow for changes to the submission platform, which has had an effect on the release of the report. The data and analysis in the report reflects information about people working within the social care workforce in the summer of 2022.

1.1   Background

This is the second time we've collected data about the entire social care workforce in Wales as a single exercise. The move to this new integrated process was agreed as part of Welsh Government's new Performance and Improvement Framework, launched in April 2020.

The social care workers included in the collection are those who provide care and support as described by the Social Services and Well-being (Wales) Act 2014.

Our aim is to capture information about the number and characteristics of everyone who makes up the social care workforce in Wales. We include information that will help us think about workforce planning, recruitment and retention. This collection is different to the data collected about the registered workforce, as it tells us about the entire social care workforce, not just those who are registered care workers.

This report includes data from local authorities and commissioned providers made up of commercial businesses, not-for-profit and third-sector organisations. It doesn't include organisations who only provide NHS-funded care, care provision paid for privately without local authority funding, or government agencies responsible for social care regulation and inspection.

We asked for the same information this year as we did last year, to allow local authorities and providers to continue to adapt their reporting systems to capture

what's needed. We asked providers of social care to submit data under the headings of:

- adult residential care

- children's residential care

- domiciliary care

- social work teams

- day services

- supported living

- personal assistants

- foster carers

- adult placement schemes

- central or support staff

- other social care services which don't fall into any of the above.

A full list of data categories included can be found in appendix A.

Introducing any new data collection framework is challenging. Some of the data we received was of insufficient quality, so we've limited the publication of data where confidence is low. For example, the number of personal assistants and foster carers in post is included in the workforce total, but not in the detailed analysis of workforce characteristics. After we talked to foster care agencies, we decided that the categories of information requested about working hours and contract type don't lend themselves well to the role of a foster carer.

How, and if, we continue to collect and represent certain information in this exercise is something to be discussed before the start of the next collection.

We calculated the number of providers we expected data from using the lists provided by local authorities, information from our internal register, and the registered services provided by Care Inspectorate Wales (CIW). Just over two thirds (68 per cent) of commissioned providers gave information about their settings and/or workforce. Although this was a slightly lower return rate than in 2021, the quality of information we received was higher. All 22 local authorities returned data to us and, overall, the completion rate for all sections of the collection has improved.

We calculated return rates this year by counting the number of organisations that provided at least a partial return (the minimum requirement being the number in post for at least one of their settings) and dividing that number by the total number of organisations we expected to return data. This is different to last year's method, as we didn't know which organisations had provided data by the time the data had been processed by local authorities and submitted to us.

We've talked to providers directly through a series of workshops and meetings, which has been a helpful learning experience for all of us. We hope to carry on with this level of engagement in the future, as we work to further improve the annual workforce data collection.

We arranged workshops to, for example, clarify the data systems and terminologies used, and to get feedback about the barriers and nuances of the data collection for organisations. This is helpful information for the work of continually developing the data collection method, system and procedures. Thanks to all providers who engaged with us during this process. Your time and knowledge is greatly appreciated.

1.2 Summary of the social care workforce in Wales

- There are an estimated 84,134 people employed in the social care workforce in Wales.

- This is a seven per cent decrease* on last year's figure.

- Of these, 31,315 (37 per cent) work in local authorities and 52,876 (63 per cent) in commissioned providers.

- 82 per cent are female and 18 per cent are male.

- Age groups remain evenly spread across the sector overall. Local authorities have an older workforce, with 52 per cent being 46 or older. Meanwhile, 56 per cent of the commissioned providers' workforce is 45 or younger.

- 95 per cent of the social care workforce was recorded as white. This is in line with the general population of Wales but represents a decrease in diversity from last year's figures (over 10 per cent non-white).

- There's been a reduction in the proportion who recorded their ethnicity as mixed (0.7 per cent of the workforce) - a reduction of four per cent on 2021 figures. This is the only ethnic group that's marginally less than the proportion found in the general population of Wales (0.3 per cent lower).

- 80 per cent of the workforce is employed on a permanent contract.

- The percentage of agency workers in the sector has decreased slightly. This is most notable in the commissioned providers, which is now 1.7 per cent (compared to three per cent in 2021). Meanwhile, the proportion of agency workers in local authorities has increased marginally to 0.5 per cent (0.4 per cent in 2021).

- 226 more people joined the sector than left this year.

- 5,323 vacancies were recorded in 2022, which equates to nine per cent of the total workforce.

See section 1.3 on revisions to calculation methods

1.3 Changes to methods for calculating missing data and data quality

We changed how we collected the data in 2022, which allowed us to understand exactly which organisations provided us with data and which didn't. This means we can be more sophisticated in our methods for dealing with missing data. But changes to these methods affected our ability to accurately compare this year's data with last year's. While comparisons have been given, it's important to remember that the change in calculation methods will likely have had some effect on the size of the variation.

While the number of providers supplying data has decreased slightly, the quality has improved significantly in some areas. This has allowed us to create more accurate estimations of the total number of people who work in social care in Wales.

There are also some anomalies within the data. For example, we calculate that there are fewer people working in social care in 2022 than in 2021, but the number of joiners is higher than leavers, so we'd expect to see a net increase rather than a decrease.

We'll continue to investigate both of these issues, and we'll publish updates about our workforce data later in the year.

While improving, data quality and missing data are still a challenge to our ability to analyse the information we receive. We'll continue to work with these data to better understand anomalies and put in place measures to reduce the impact of errors and improve our collection.

There's more detailed information about our steps for dealing with missing data in appendix B.

1.4 Lessons learned and next steps

This is the second year we've carried out this data collection and while we've identified some new issues, there are many positives.

Those who submitted data provided us with higher-quality submissions which were more complete and more accurate. This is thanks to the efforts of those who've collected and returned data and the data validation routines put in place this year.

Social care providers who engaged with us did so in a positive and constructive way which, given the pressures in health and social care in Wales, was gratefully appreciated.

We've identified the need for three key changes going forward. We'll:

- review and change how commissioned provider settings are recorded in our system to make sure the associated contact details for settings and organisations reflect work e-mails rather than personal accounts

- review the information collected for personal assistants and focus on how this can be made easier to collect and the quality improved

- reassess the information requested about foster carers and whether this exercise is the best and most appropriate method for capturing this data.

We'll continue to work with other organisations to make sure the collection remains relevant and comparable with other data sources.

We discovered during our data validation checks that there are large differences in the number of settings and individuals employed in the three residential categories (children's, adult and mental health) this year compared to last year. But the combined figures for all three residential categories over the two years are similar.

	Total number of settings	Total number in post
2021	976	34,543
2022	978	34,342

Table showing residential care comparison between 2021 and 2022.

We assume that these differences are because services have been categorised differently. It's possible that commissioned providers have chosen a different

category to those previously chosen by the local authority. This is something we're keen to explore over the course of the next year.

	2021		2022	
	No. in post	No. of settings	No. in post	No. of settings
Adult - residential care	30,531	744	29,100	757
Children - residential care	1,838	107	4,411	191
Mental health - residential care	2,174	125	831	30
Total	34,543	976	34,342	978

Table showing difference in settings for residential care between 2021 and 2022.

2. Conclusion

Introducing a large, new data collection activity is challenging and can take time to become a consistently reliable process. There are still some limitations in the data collected in 2022 and there have been issues with our publication timescales. We'll continue to work with our partners to improve the process year on year.

We've made progress in several key areas this year. We improved our collection methods, our validation process and, despite having a slightly lower completion rate, we've received higher-quality data than in 2021. We've also improved our models for imputing missing values to determine our overall estimates for the total number of people working in social care in Wales.

We've also been working with Welsh Government to make sure the annual workforce data collections are put on a more solid statutory footing, which should help to improve the data we're able to collect for this important work.

As this is the second year of collection, we can draw some comparisons with last year. This allows us to more easily identify areas where further discussion and investigation is needed around differences in data.

As we complete more collection cycles, we must explain these differences to develop a true understanding of what's happening in the sector.

We're looking forward to building on the relationships established this year, and supporting those who carry out the collection to be able to provide their data in a meaningful way.

3. Our findings

3.1 Overall social care workforce

We received data for more than 61,000 people working in social care. We received a 100 per cent return rate from local authorities, but information from commissioned providers was incomplete at only 57.9 per cent.

Local Authority

Commissioned Provider

Returns (%)   Not Returned (%)

n = 22   n = 869
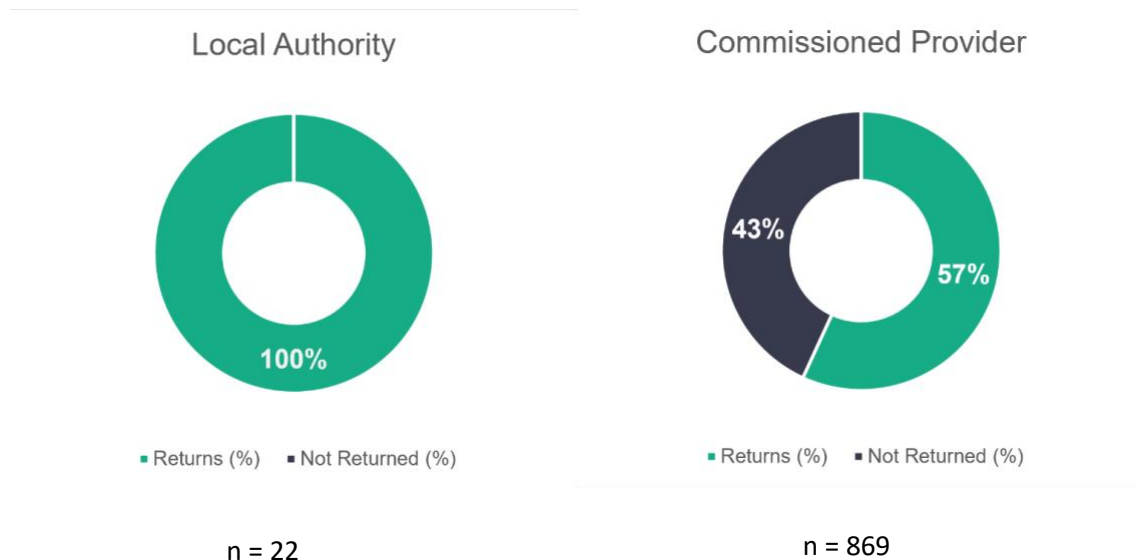
Using data modelling techniques (explained in more detail in appendix B), we created an estimate for the total number of people working in social care in Wales. We estimate this figure to be **84,134**.

While our estimates show a lower estimate when compared to our 2021 report (91,000), it's important to take into account the change in the way we calculated missing data this year.

Personal assistant and foster carer numbers are included in the above figure but aren't included in the following charts.

We estimate that the workforce is split across service type as in the following chart:

Social care workforce by service area
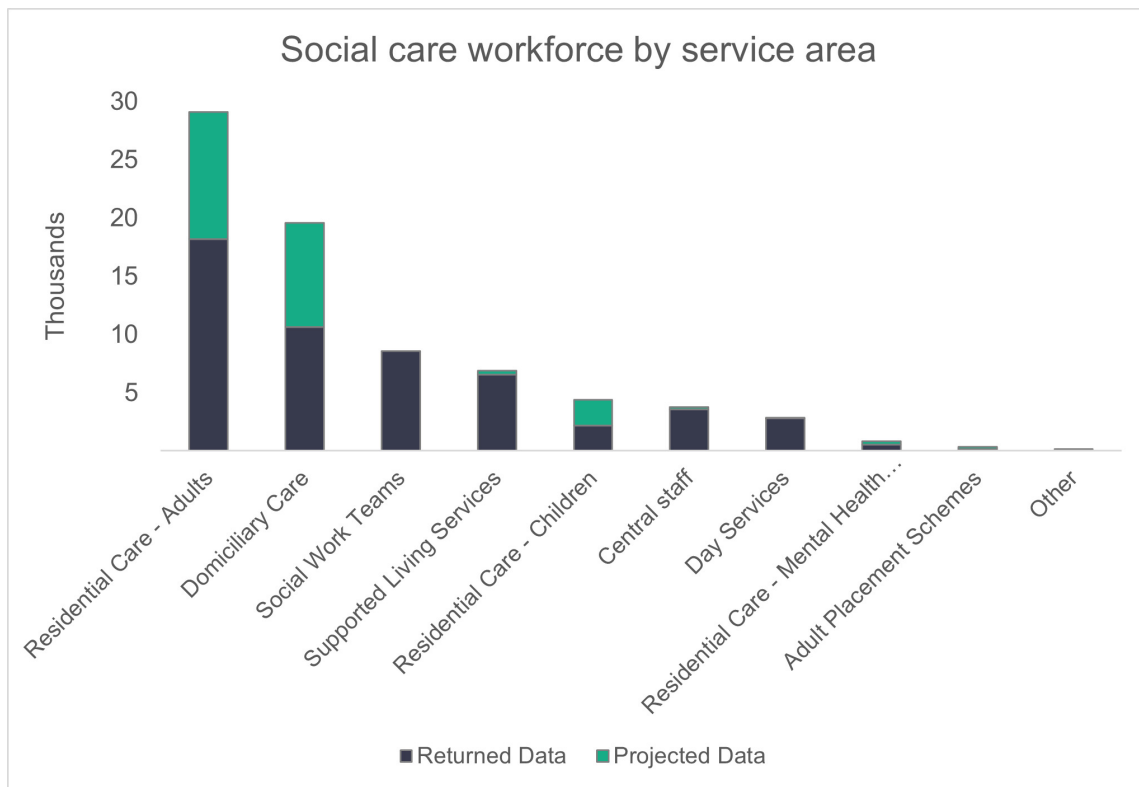
n = 76,568

3.1.1 Gender

More than four in every five social care workers in Wales are women. Men make up 18 per cent of the total workforce. A total of 41 workers were recorded as gender fluid or non-binary, which represented 0.08 per cent of the submitted data (preferred not to say responses aren't included).
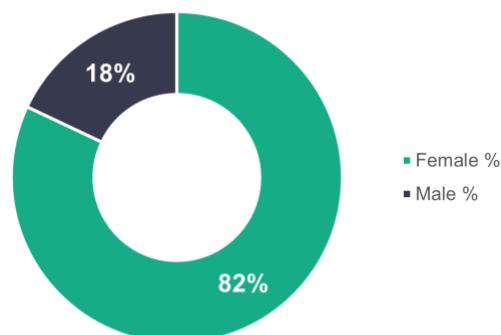
n = 46,261

3.1.2 Age profile

The chart below shows the age distribution for the social care workforce in Wales and is expressed as a percentage of all local authority-employed workers and percentage of all commissioned services. The workforce is distributed through the whole working life range and tapers off at the early and late (16 to 25 and over 65 groups) career ages. Commissioned service providers have a generally younger workforce compared to local authority providers, with more workers aged 26 to 45 for commissioned services and more workers in the 46 to 65 range for local authorities.
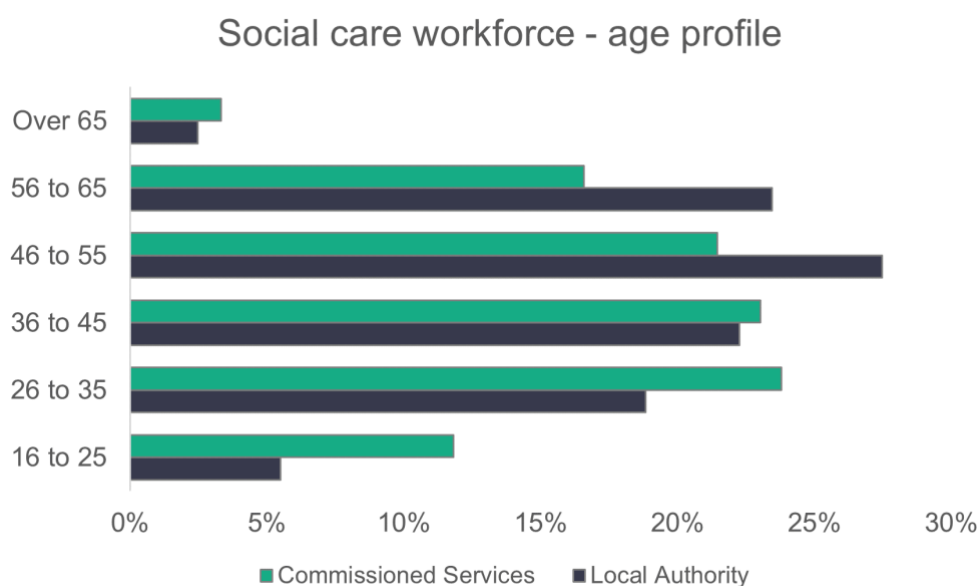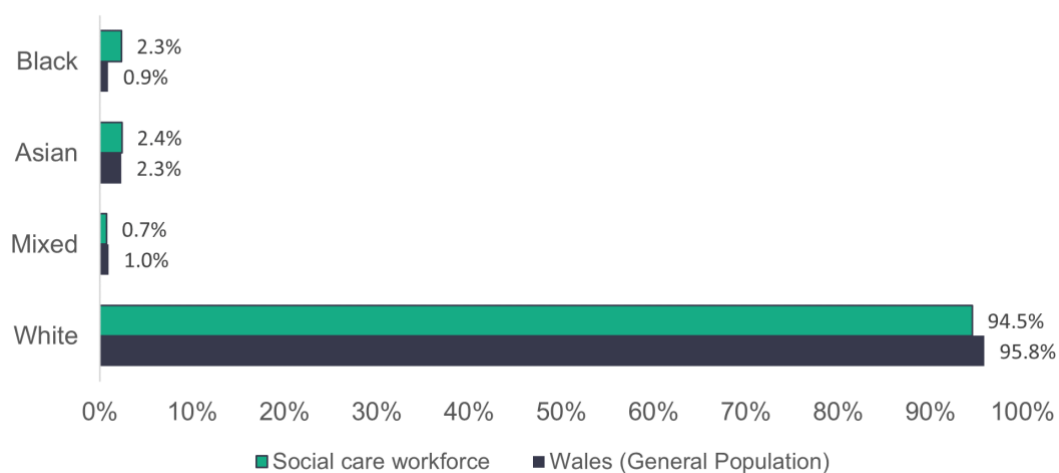

n = 45,494

3.1.3 Ethnicity

The ethnicity of the social care workforce broadly mirrors that of the Welsh population. The exception is in the proportion of black (or black British) workers in the social care workforce, which is around 2.5 times more than in the Welsh population, when viewed proportionally.

In comparison to the 2021 data, the social care workforce is now less ethnically diverse, with 94.5 per cent reported as white in 2022, compared to 89.1 per cent in the previous year.
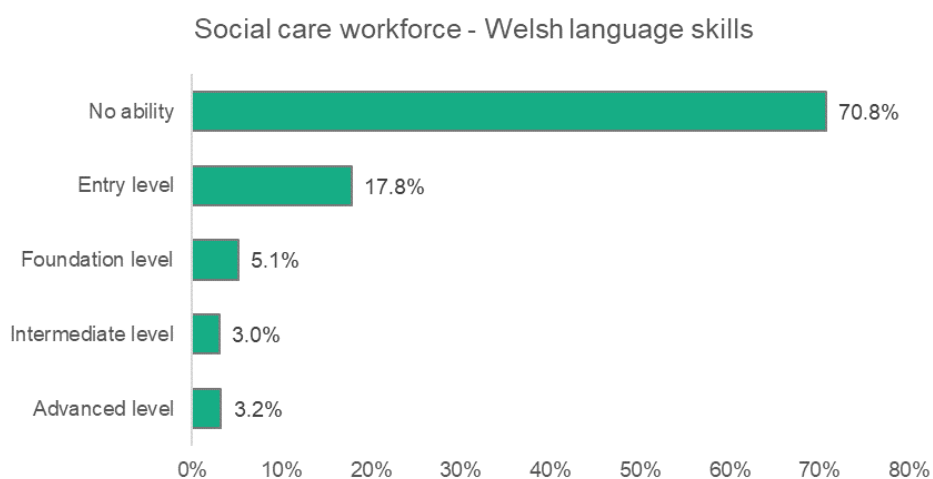
Social care workforce - ethnicity

Black	2.3% / 0.9%
Asian	2.4% / 2.3%
Mixed	0.7% / 1.0%
White	94.5% / 95.8%

Legend: ■ Social care workforce   ■ Wales (General Population)

n = 33,362

3.1.4 Welsh language skills

Our data suggests that around 29 per cent of the social care workforce in Wales can understand Welsh to some degree. A breakdown of their Welsh language proficiency is shown below.

.

Social care workforce - Welsh language skills

|	
No ability	70.8%
Entry level	17.8%
Foundation level	5.1%
Intermediate level	3.0%
Advanced level	3.2%

Data reported by Welsh Government suggests that just over 22 per cent* of the population aged over three in Wales can speak Welsh, with two per cent able to speak a few words and 11 per cent fluent Welsh speakers.

Data collected from the Annual Population Survey in 2021 states that 29.1 per cent* of the population in Wales can speak Welsh.

Both of these data use slightly different categorisation to the Welsh Language Skills framework used in the workforce data collection. These are comparable to the Common European Framework of Reference for Languages (CEFR) and the Association of Language Testers in Europe (ALTE) frameworks used elsewhere.

3.1.5 Contract type

Overall, 80 per cent of those employed in social care in Wales are employed on permanent contracts, a slight increase from last year's 78.5 per cent. We only found a slight difference between the proportion of those on permanent contracts for local authority-run services and commissioned services.

Social care workforce - contract type
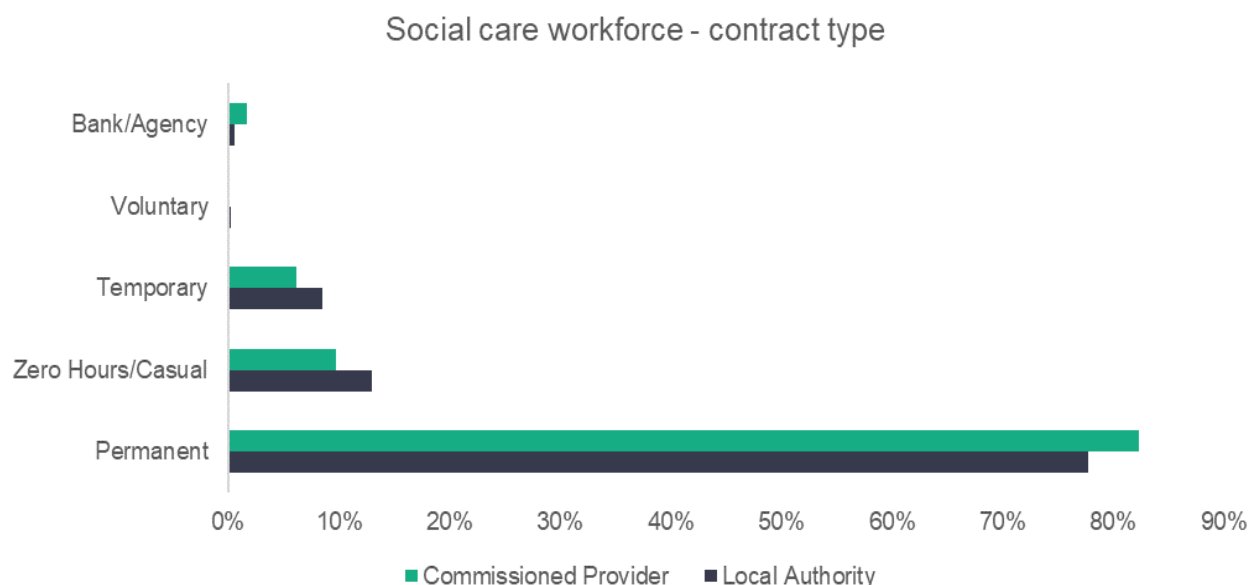
n = 47,317

The data includes workers on casual and/or zero hours contracts and shows that local authority-run services employ a higher proportion of their staff on these contracts (12.9 per cent for local authorities compared to 9.7 per cent for commissioned services). This is different from last year's figures, when 16.5 per

cent of staff in commissioned services and 11.7 per cent of staff in local authorities were on casual and/or zero hours contracts.

But we recognise that casual and zero-hours contracts are very different and our current collection methods aren't able to distinguish between them. We'll look to separate these for the 2023 data collection to improve our intelligence.

Commissioned services employ slightly more bank or agency staff in relation to their entire workforce (1.7 per cent for commissioned services, compared to 0.5 per cent for local authorities). The proportion of bank or agency staff for commissioned services decreased by approximately 1.3 per cent compared to last year, while the proportion for local authorities increased by 0.1 per cent.

3.1.6 Working patterns

Working patterns in local authorities and commissioned services are broadly similar for full-time employment (36 hours or more). There's a marked difference in the number of people working up to 16 hours a week between commissioned services and local authorities.

Social care workforce - working patterns (hours per week)

n = 46,270
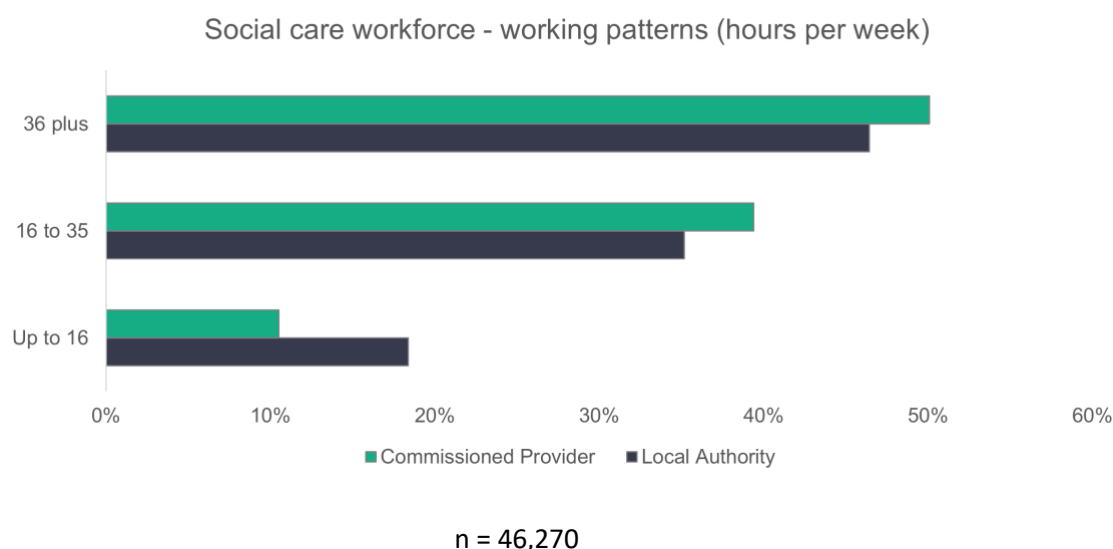
3.1.7 Vacancies

The number of vacancies reported in the workforce data collection process was 4,257. An extra 730 posts were being 'held vacant' by employers. Using the same methods for estimating the total social care workforce in Wales, we can estimate an additional 1,066 posts that are vacant or held vacant in the sector, giving a total

of 5,323 (+ /- five per cent) posts vacant or held. This represents nine per cent of the total workforce.

The number of actual vacancies reported has increased this year by 1,581 (2,676 recorded in 2021), with a reduction of 945 (1,675 in 2021) in the number of vacancies being held, resulting in an overall increase of 636 total vacancies this year (total vacancies in 2021 was 4,351, and 4,987 in 2022).

Two-fifths (38.6 per cent) of all vacancies are believed to be in local authorities, an increase of 6.2 per cent on the 2021 figure. Commissioned providers account for 61.4 per cent of the vacancies, a decrease of six per cent on the 2021 figure.

Social care workforce - vacancies

Setting	Vacant	Held	Estimated data
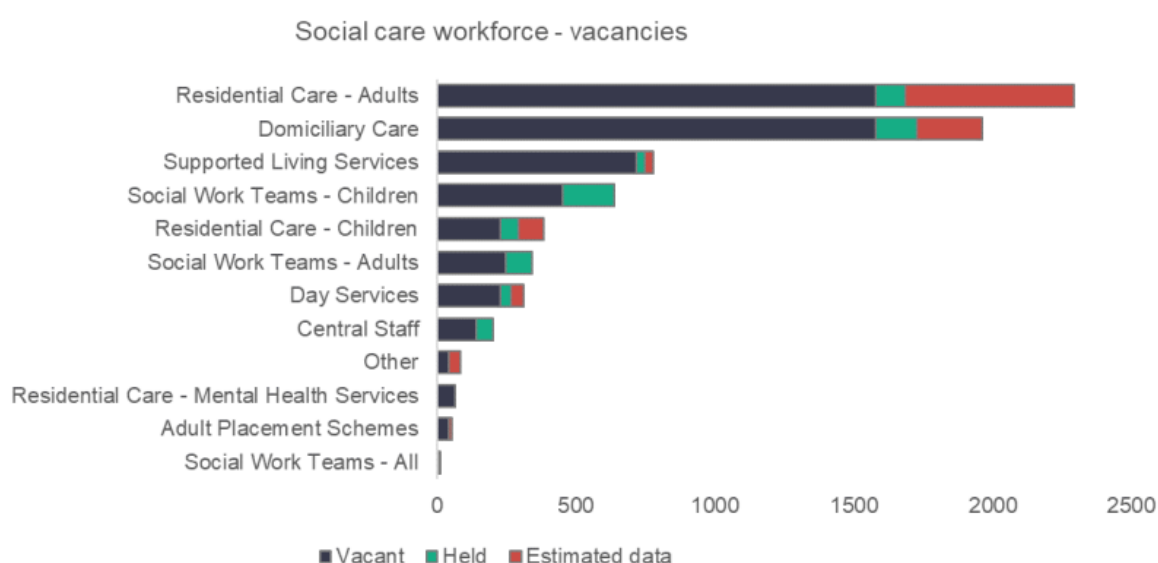
3.1.8 Staff turnover

We asked for information about starters and leavers. It was reported that 1,593 people started a new role in social care in Wales during the year, and 1,367 left social care. From these data we calculate an overall net increase of 226 staff in the social care sector in Wales from 2021 to 2022. This data doesn't correlate with our expectations from our estimation of the total number of staff working in social care in Wales.

Setting	Joiners	Leavers	Net change
Adult placement schemes	15	10	+5
Central staff	48	18	+30

Day services	22	32	-10
Domiciliary care	53	68	-15
Residential care - adults	61	56	+5
Residential care - children	496	372	+124
Social work teams - children	802	686	+116
Supported living services	96	125	-29
Total	1,593	1,367	+226

Table showing joiners and leavers between 2021 and 2022.

Our estimations suggest a decrease in the number of workers in social care in Wales while the turnover figures suggest a marginal increase. This difference may reflect the changes to the method of collection, but it may also be influenced by our improvements to the method for estimating the missing data.
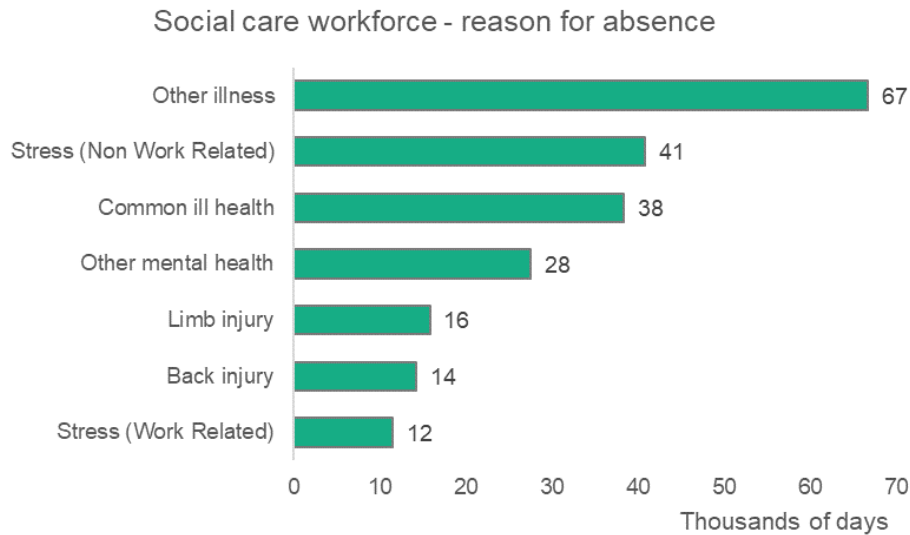
Comparisons between 2021 and 2022 give a valuable insight and we acknowledge that while our new methods give a lower figure, they're also likely to be more accurate than our estimates in 2021. Our new approach to collecting data, as used in 2022, will now form the basis of future collections.

We're also aware of continued pressures in social care and that the data collected, particularly those data about vacancies, match the anecdotal evidence in social care in Wales.

3.1.9 Sickness

We asked for information about sickness in the workforce data collection process. Completion of these data in 2022 was mandatory for local authorities but optional for commissioned services. The data we received showed 214,941 days lost to sickness. While this is almost certainly an underestimate given the data we received was incomplete, it's a significant increase on last year (85,848 days reported lost).

The reasons for these absence days remains relatively consistent, with other illness being the highest at 31.0 per cent (36.9 per cent in 2021). This is followed by stress (non-work related) at 18.9 per cent (15.6 in 2021) and common ill health at 17.6 per cent (14.3 per cent in 2021). Seventy per cent of the days were recorded as lost to long-term sickness, a six per cent decrease from 2021, and 30 per cent to short-term, which is an increase of six per cent from 2021.

Social care workforce - reason for absence

| Reason	Thousands of days
Other illness	67
Stress (Non Work Related)	41
Common ill health	38
Other mental health	28
Limb injury	16
Back injury	14
Stress (Work Related)	12

Other illness was the dominant single category, accounting for 31.0 per cent of all reported sickness absence. But if we consider mental health as a whole, (combining both stress categories and other mental health) it becomes the largest grouping, accounting for 37.2 per cent of sickness absence.

We'll continue to work with data collectors to make sure they understand what each of these categories includes, and to help keep our reporting systems and our data categorisation in line, wherever possible.
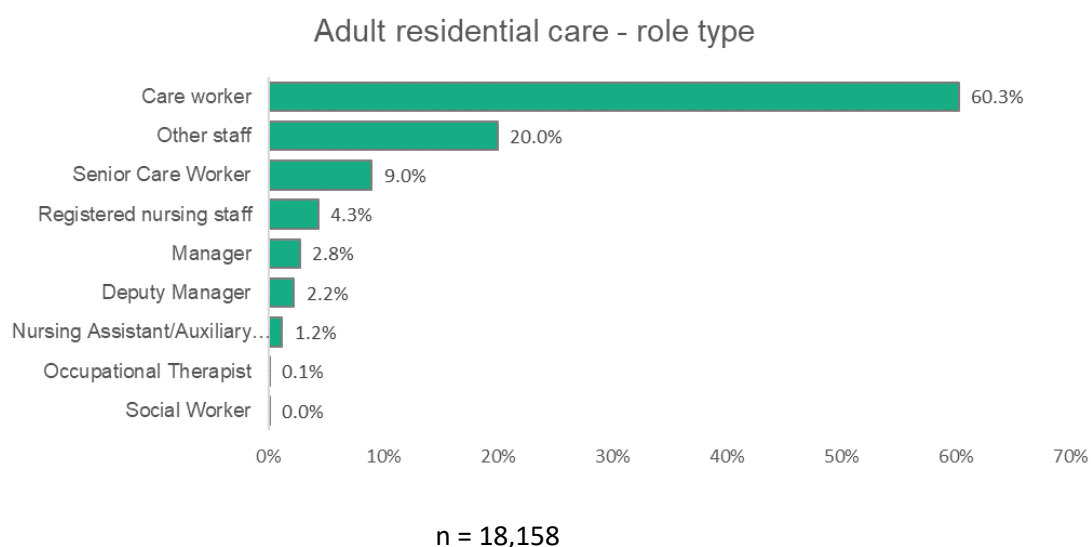
3.2 Adult residential care

We estimate the total number of people working in adult residential care in Wales in 2022 to be 29,100. In 2021 we estimated the figure to be 30,531 - a drop of 4.6 per cent in the estimated workforce over the last 12 months.

Return rates for local authority-run services were much higher than those from commissioned services, as we see below.

n = 18   n = 501
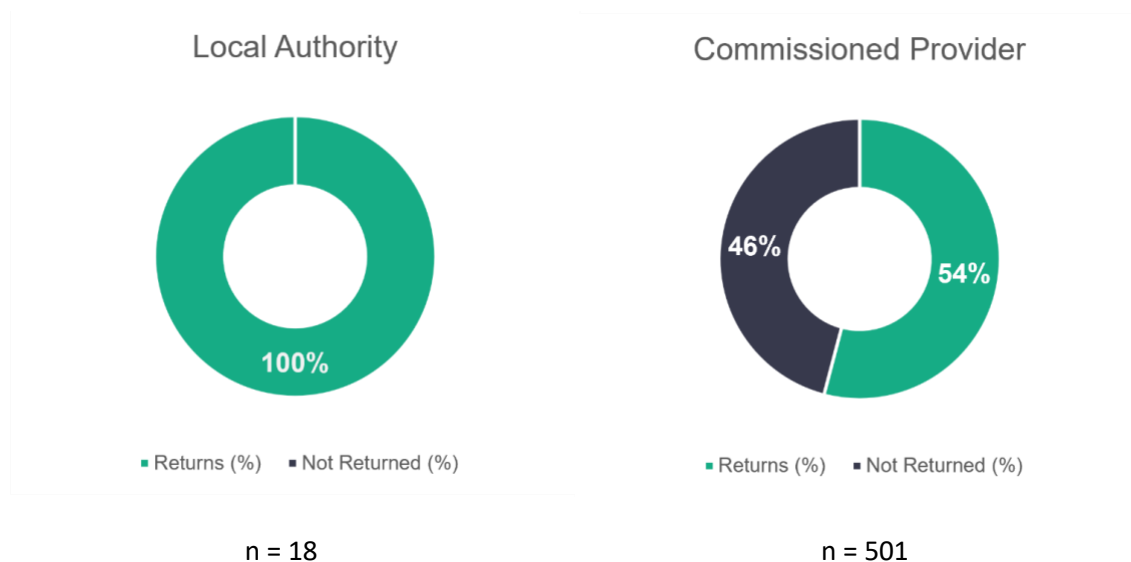
3.2.1 Adult residential care – role type

Care workers make up 60.3 per cent of the adult residential care workforce, as shown below. The next largest category is other and is likely to include ancillary and business support staff.

n = 18,158

3.2.2 Adult residential care – gender

Women occupy slightly more adult residential care roles than we'd see across the whole social care workforce in Wales. Fewer than five non-binary and gender-fluid people were recorded. These numbers were too small to be depicted in the chart below.
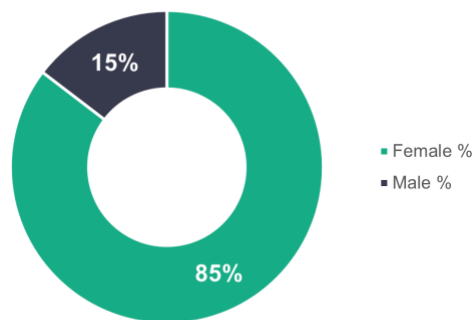
n = 14,023

### 3.2.3 Adult residential care – age profile

The age profile for adult residential care workers is broadly similar to the general social care workforce. A higher proportion of those working in local authority adult residential care are in older age groups, while workers in commissioned adult residential care are more equally spread out across the age groups. Similar proportions (3.3 per cent) are in the over 65 years age group in local authorities and commissioned services.

Adult residential care - age profile
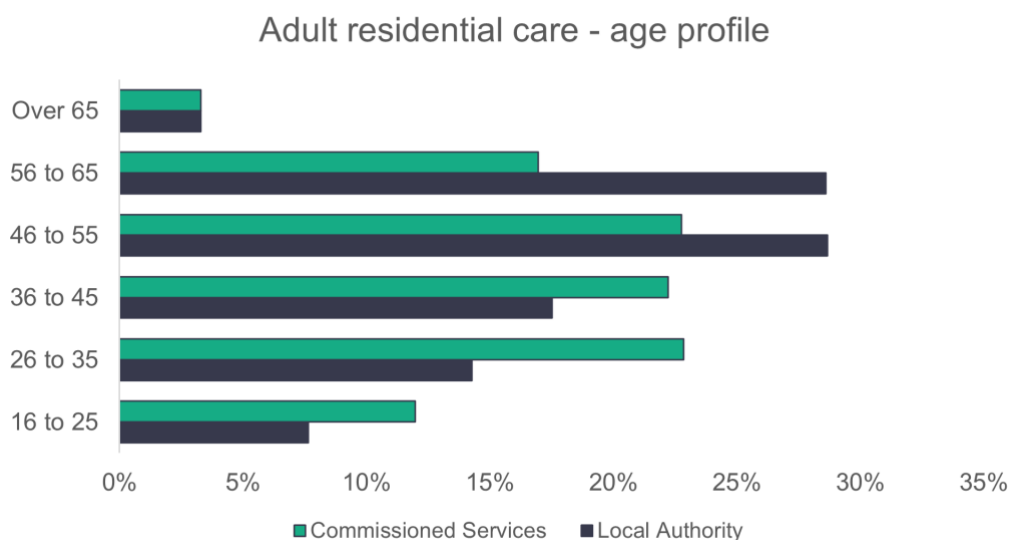
n = 13,258

3.2.4 Adult residential care – ethnicity

Local authorities and providers reported that the proportion of black or black British, and Asian or Asian British workers is higher in adult residential care than in the general population of Wales. But compared to last year, the diversity of the adult residential care workforce has fallen (2021: 3.0 per cent black, 7.8 per cent Asian and 1.1 per cent mixed).
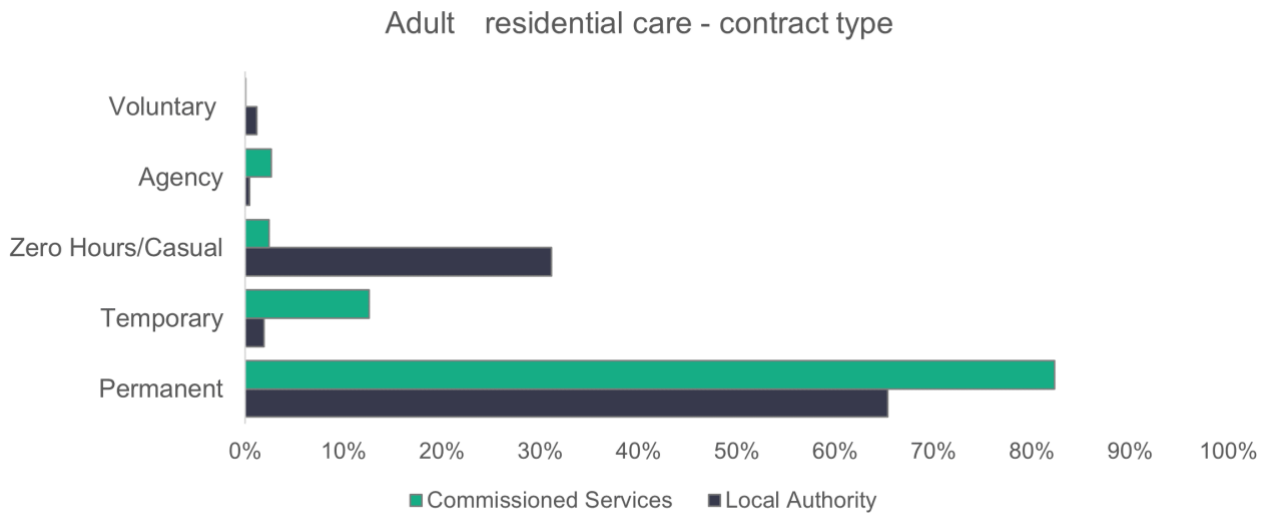
Adult residential care - ethnicity

Ethnicity	Adult residential care	Wales (General Population)
Black	2.4%	0.9%
Asian	5.9%	2.3%
Mixed	0.5%	1.0%
White	91.2%	95.8%
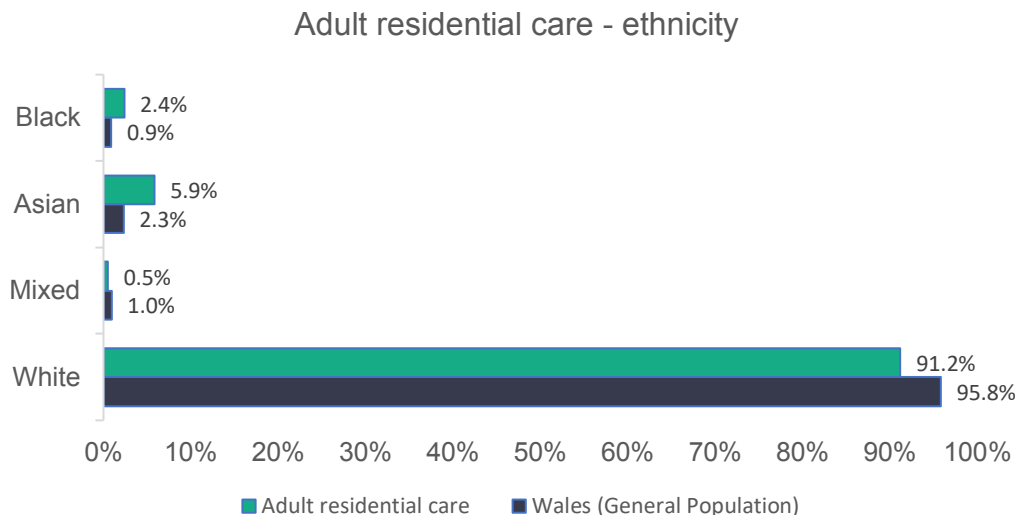
n = 9,082

3.2.5 Adult residential care – contract type

Permanent contracts account for 82.6 per cent of staff employed in commissioned adult residential care services, a reduction of 7.4 per cent on the previous year's figure (90 per cent in 2021). A further 12.4 per cent are employed on temporary contracts. We found that 2.4 per cent of people who work in commissioned adult residential care are employed on a casual or zero-hours contract.

The picture for local authority adult residential care services looks slightly different, with only 66.1 per cent of these staff employed on a permanent basis, a reduction of 3.2 per cent on the 2021 figure. Meanwhile, 31.5 per cent are employed on a casual or zero hours contract, an increase of around six per cent on the 2021 figure.

Adult   residential care - contract type

n = 14,715

3.2.6 Adult residential care – working patterns

More than a third (37 per cent) of those working in local authority adult residential care work less than 16 hours per week, a slight increase from the 33.2 per cent that was recorded last year. Also, 8.3 per cent of the commissioned adult residential care workforce work less than 16 hours per week, which is a minor decline from the figure of around 10 per cent that we noticed last year. Over half (51.5 per cent) of those in commissioned adult residential care work full-time hours, a small rise from the 48.7 per cent that we observed last year.
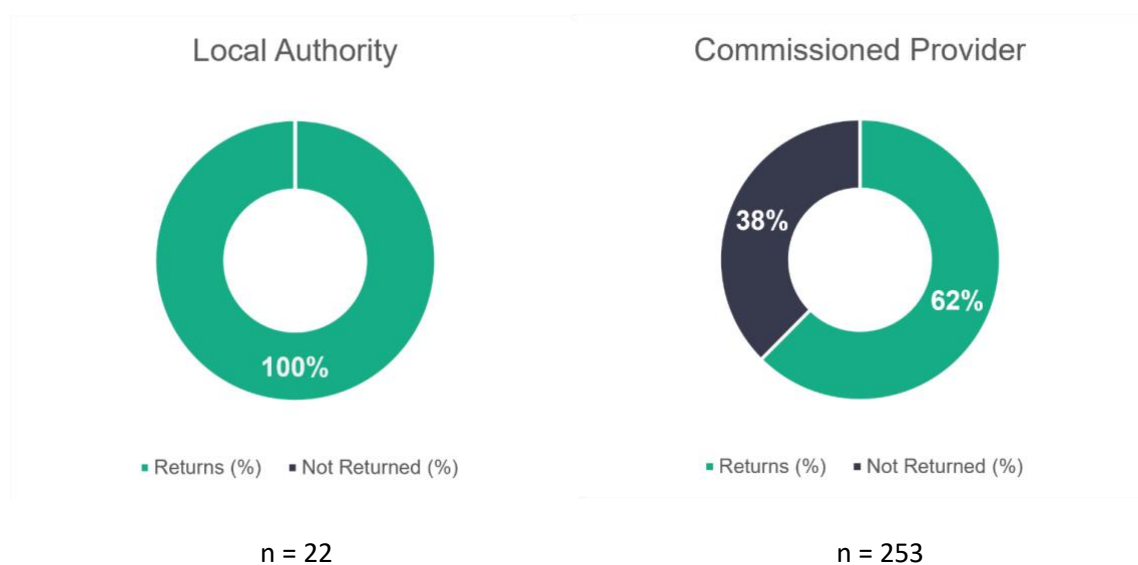
Adult residential care - working patterns (hours per week)
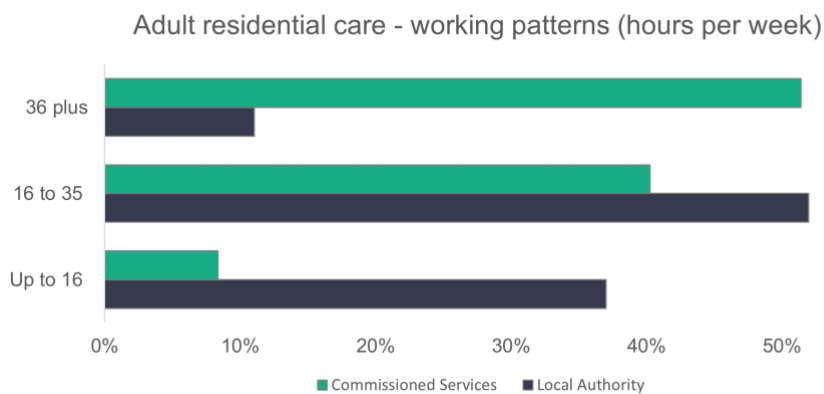
n = 14,348

3.3 Domiciliary care

We estimate that the total number of people working in domiciliary care is 19,571.

Of those, 4,232 (21.6 per cent) come from local authority-run services and 15,339 (79.4 per cent) come from commissioned service providers. In 2021, we estimated 23,108 people worked in domiciliary care, which shows a drop of 18.1 per cent in the workforce.
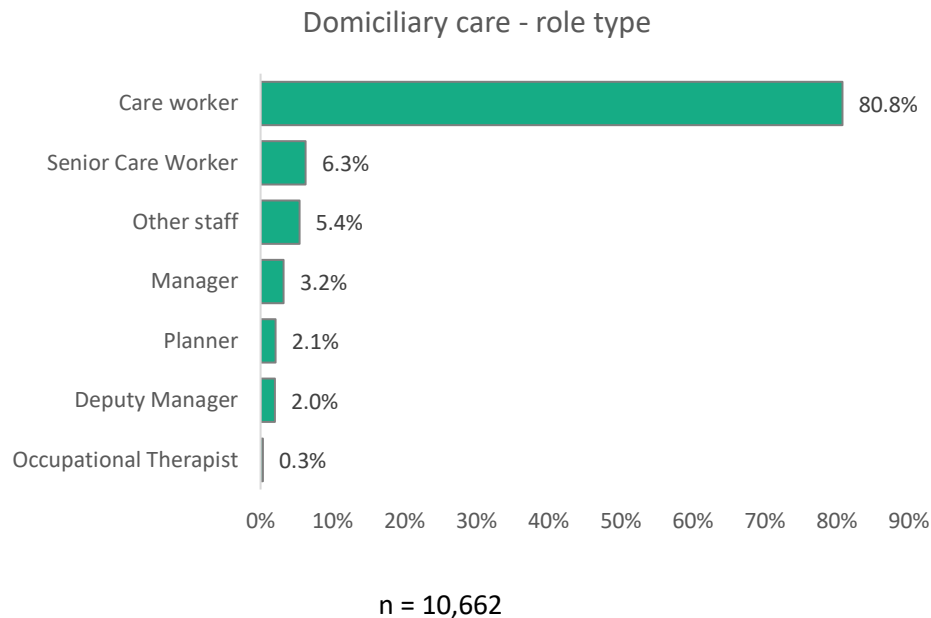
We must be cautious when comparing with last year. While a fall in the number of workers corresponds with anecdotal evidence from social care in Wales, we can't determine how much of this difference is because of our change in the calculation method.

Local Authority — Returns (%) ■ Not Returned (%) — 100%

n = 22

Commissioned Provider — Returns (%) ■ Not Returned (%) — 38% Not Returned, 62% Returns

n = 253

The return rate for the number of settings this year is 62.5 per cent for commissioned services, which is slightly lower than the 66.8 per cent returned last year. Local authority services had a return rate of 100 per cent again this year.

3.3.1 Domiciliary care – role type

Care worker roles make up most staff in this service area, as expected, accounting for 80.8 per cent.

Domiciliary care - role type

| Role	Percentage
Care worker	80.8%
Senior Care Worker	6.3%
Other staff	5.4%
Manager	3.2%
Planner	2.1%
Deputy Manager	2.0%
Occupational Therapist	0.3%

n = 10,662

3.3.2 Domiciliary care – gender

The proportion of female to male workers is slightly higher than that reported for the social care sector as a whole. We found that 88.1 per cent of domiciliary care roles were occupied by women, compared to 81 per cent of women in the overall workforce. Fewer than five domiciliary care workers were reported as being gender fluid and fewer than five were reported as non-binary. A total of 346 preferred not to state their gender.
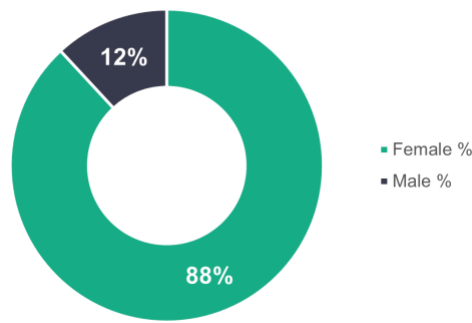
n = 9,095

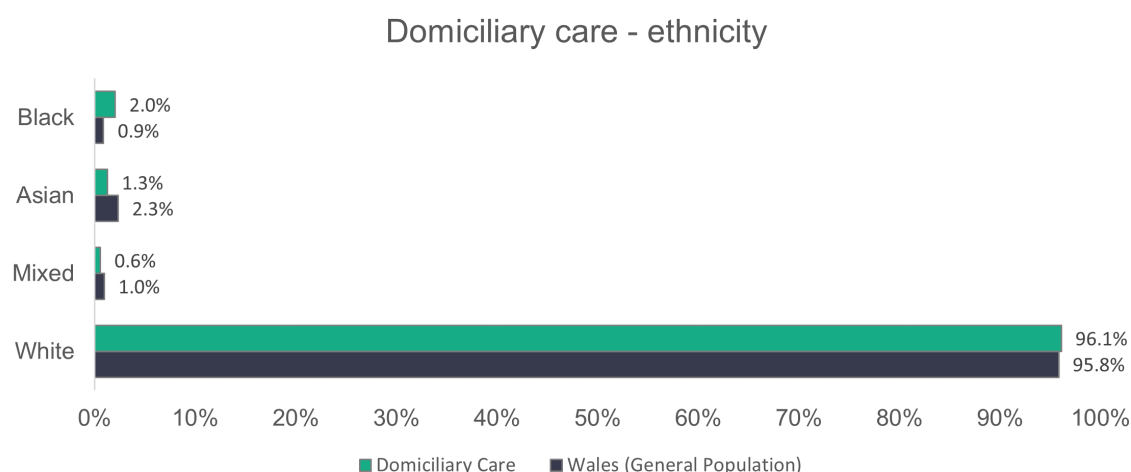### 3.3.3 Domiciliary care – age profile

There's a marked difference between the age profile of those working in local authority and commissioned services. Workers in local authorities are more likely to be older than those working in commissioned services, with local authority staff dominating the 46 to 65 age groups. Commissioned services have significantly higher workforce populations in younger age groups. There could be several reasons for this, including differences in terms and conditions, remuneration and turnover rates.

Domiciliary care - age profile
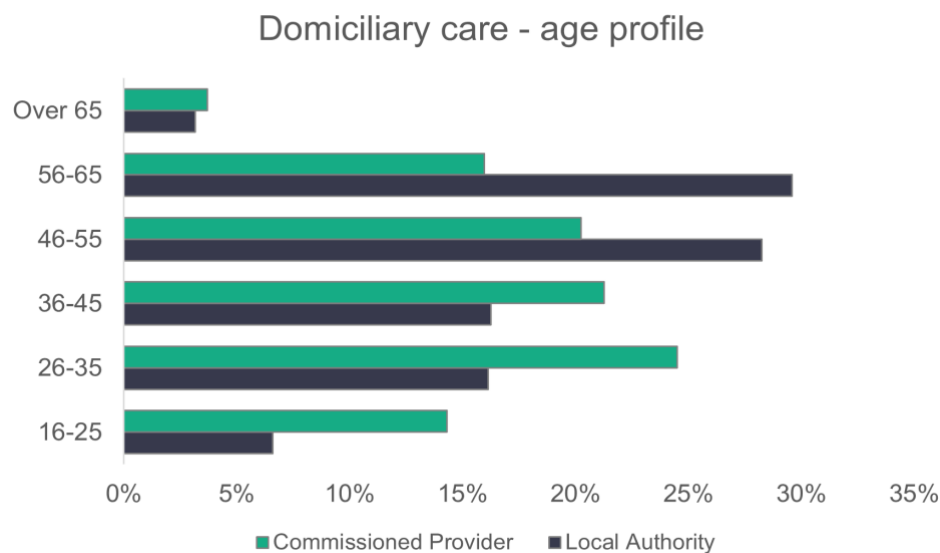
n = 9,057

3.3.4 Domiciliary care – ethnicity

Though the ethnic diversity of domiciliary care workers resembles the Welsh population, the small variations indicate a slightly less ethnically diverse workforce, particularly within black and Asian groups. The corresponding data in 2021 showed a more diverse domiciliary care workforce, with a lower proportion of staff reported as white (87.3 per cent in 2021 versus 96.1 per cent in 2022). It's also notable that those reported as 'mixed ethnicity' has fallen from 6.1 per cent in 2021 to 0.6 per cent in 2022.

Domiciliary care - ethnicity

Ethnicity	Domiciliary Care	Wales (General Population)
Black	2.0%	0.9%
Asian	1.3%	2.3%
Mixed	0.6%	1.0%
White	96.1%	95.8%

n = 6,932

3.3.5 Domiciliary care – contract type

Most domiciliary care workers are employed on permanent contracts. Local authority employers have a higher proportion of their workers on permanent contracts, at 85.1 per cent. The corresponding proportion employed by commissioned providers is 64.4 per cent.

Casual or zero-hours contracts account for 33.74 per cent of the commissioned provider domiciliary care workforce and 11.2 per cent of the local authority workforce. Agency workers make up a tiny fraction (0.5 per cent) of the total domiciliary care workforce in Wales, suggesting either a lack of availability of agency staff or better opportunities for more suitable permanent positions with providers.

There's been a shift away from zero-hours and casual employment in commissioned services (down 22.8 per cent). Many of these now seem to be employed on a permanent basis, with this category seeing a 20.8 per cent increase in 2022. Local authorities have experienced little change in the distribution of contract types, with only a slight decrease of 2.5 per cent of workers on permanent contracts.

Domiciliary care - contract type
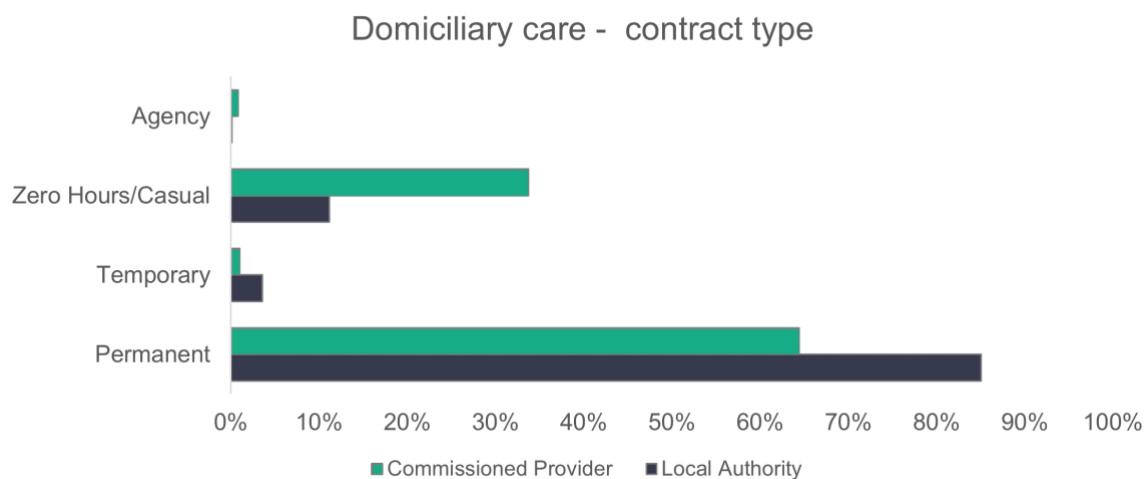
n = 9,180

3.3.6 Domiciliary care – working patterns

Most domiciliary care workers work between 16 and 35 hours per week, with a higher proportion (60 per cent) of local authority employed workers falling into this bracket. The corresponding proportion for commissioned care providers is 44 per cent, which is only slightly higher than the 38 per cent who work 36 hours or more per week.

Domiciliary care - working patterns (hours per week)
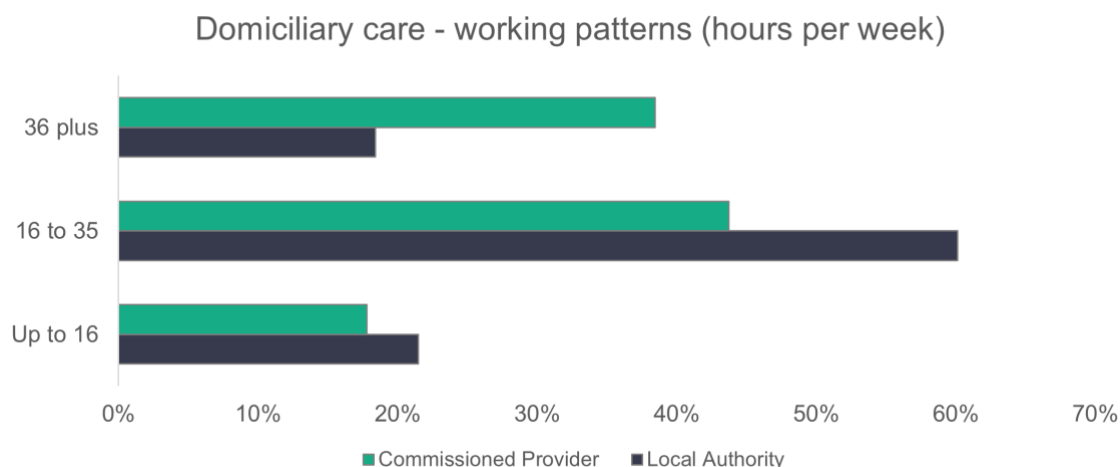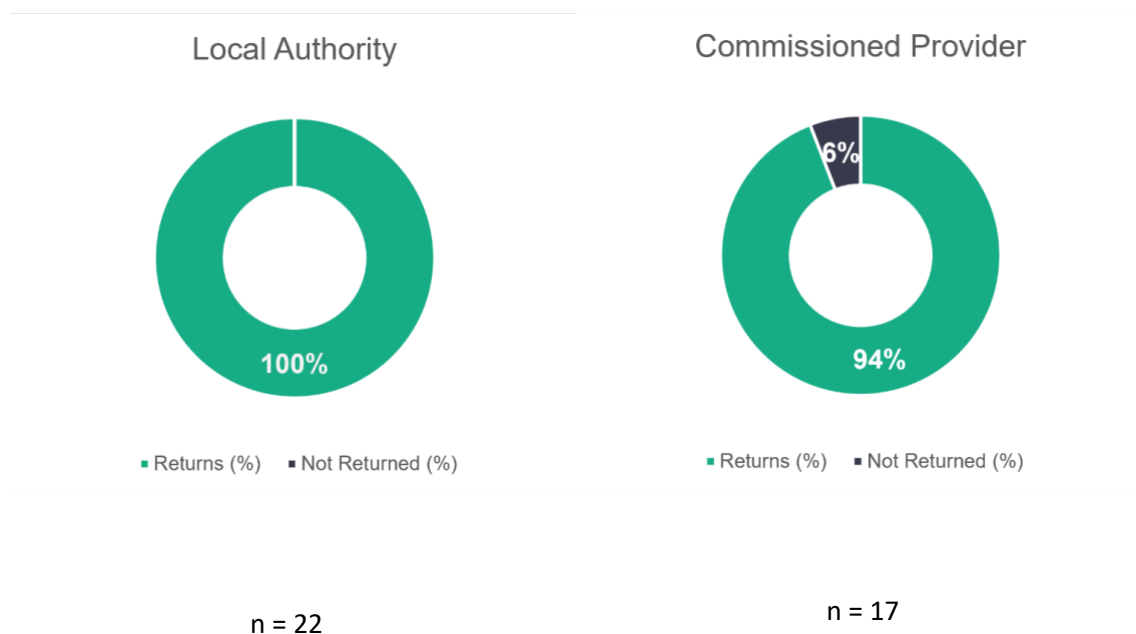

n = 8,851

3.3.7 Domiciliary care – vacancies

Domiciliary care roles account for the largest number of vacancies in the social care sector in Wales, despite being the second-largest service. There were 1,492 reported (and an additional 236 unreported but predicted) vacancies in domiciliary care that make up 29.1 per cent of the total vacancies in social care in Wales.
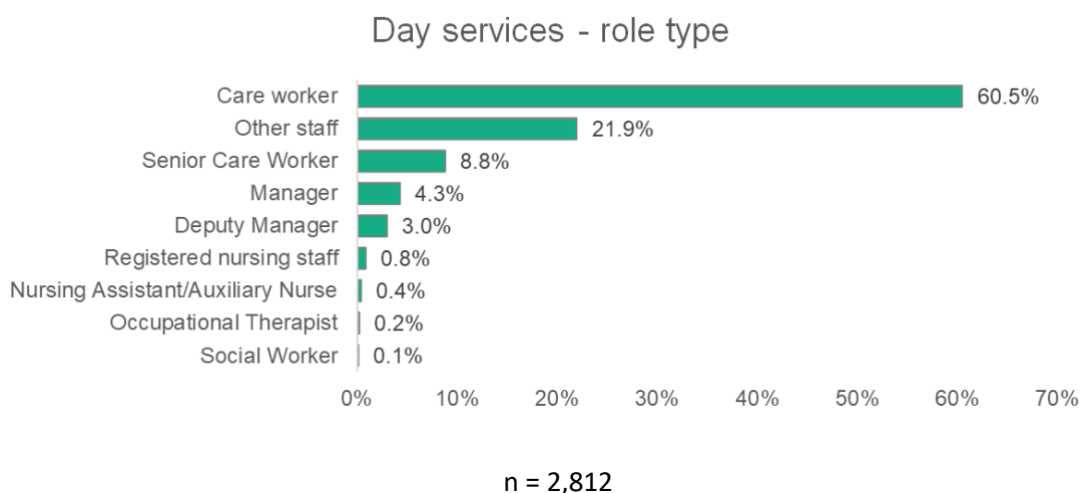
3.4 Day services

The number of workers in day services is estimated to be 2,954 (2,157 in local authority run services and 797 in commissioned services) and shows a slight increase on the 2021 workforce of 2,857.

Return rates for commissioned services were much higher this year, at 94.1 per cent compared to 61 per cent last year. Local authority service providers had a 100 per cent return rate again this year.

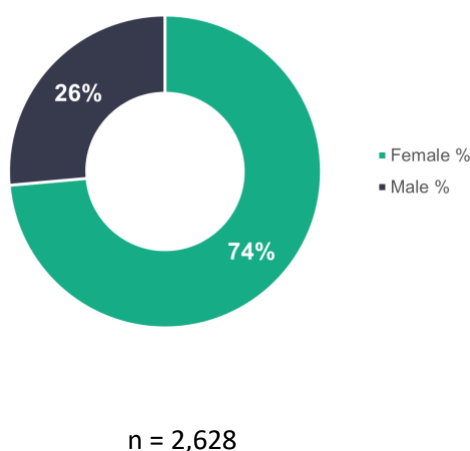| Local Authority	Commissioned Provider
100%	6% / 94%
■ Returns (%) ■ Not Returned (%)	■ Returns (%) ■ Not Returned (%)
n = 22	n = 17

3.4.1 Day services – role type

Care workers make up 60.5 per cent of the day services workforce, as shown below. The next-largest category is 'other staff', which is largely made up of business support and auxiliary staff.

Day services - role type

| Role	%
Care worker	60.5%
Other staff	21.9%
Senior Care Worker	8.8%
Manager	4.3%
Deputy Manager	3.0%
Registered nursing staff	0.8%
Nursing Assistant/Auxiliary Nurse	0.4%
Occupational Therapist	0.2%
Social Worker	0.1%

n = 2,812

3.4.2 Day services – gender

In line with most social care professions in Wales, more women than men work in day services. But the percentage of men is higher than we expect to see across all of social care, at 26 per cent (compared to 19 per cent overall).

- Female %: 74%
- Male %: 26%

n = 2,628

3.4.3 Day services – age profile

There's a distinct separation of age groups between commissioned services and local authority. Commissioned services are dominated by a younger, 26 to 45 age group (which makes up 52 per cent), and local authority by an older, 46 to 65 age group (which makes up 58 per cent).


Day services - age profile
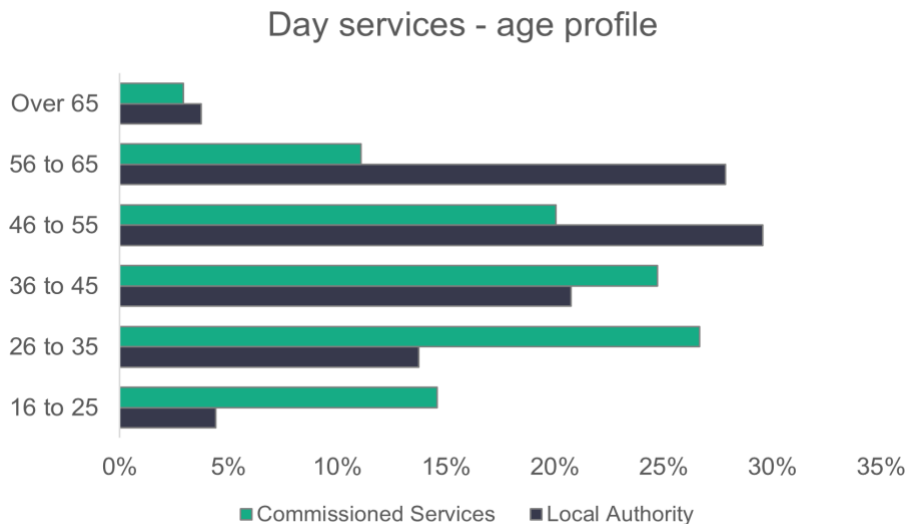
n = 2,626

3.4.4 Day services – ethnicity

The ethnicity of the day service workforce and the Welsh population are quite similar. There's a marked difference from corresponding data in 2021, which showed a more ethnically diverse day services workforce, where almost one in five workers were reported to be from a mixed ethnic background.

Day services - ethnicity
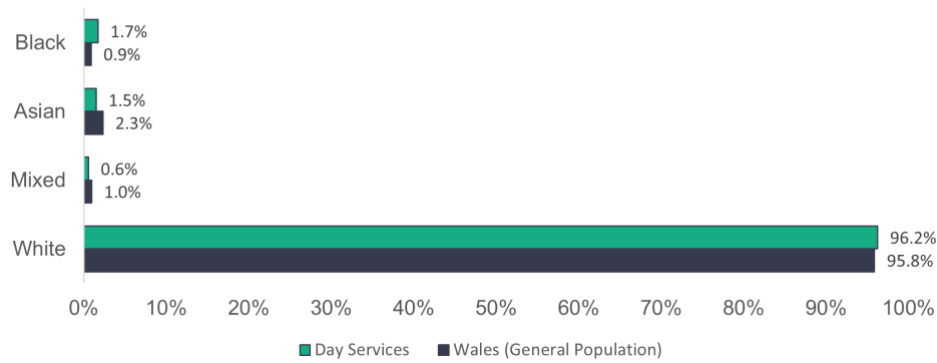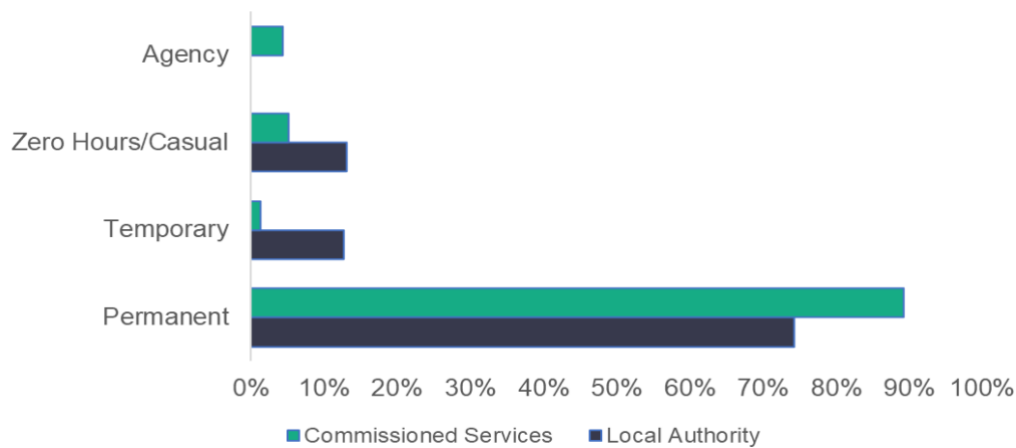
n = 1,798

3.4.5 Day services – contract type

Most staff are employed on permanent contracts for both commissioned services and local authority, with as many as 90 per cent of commissioned services employees on permanent contracts. Over twice as many day service workers in local authorities are employed on a zero-hours or casual contract.


Day services - contract type

n = 2,665

3.4.6 Day services – working patterns

Most employees in both local authority and commissioned services work full-time hours, with 50 per cent of commissioned service employees working 36 hours or more per week. Compared to those working in commissioned services, a much smaller proportion of staff in local authorities work less than 16 hours.

Day services - working patterns (hours per week)

n = 2,654
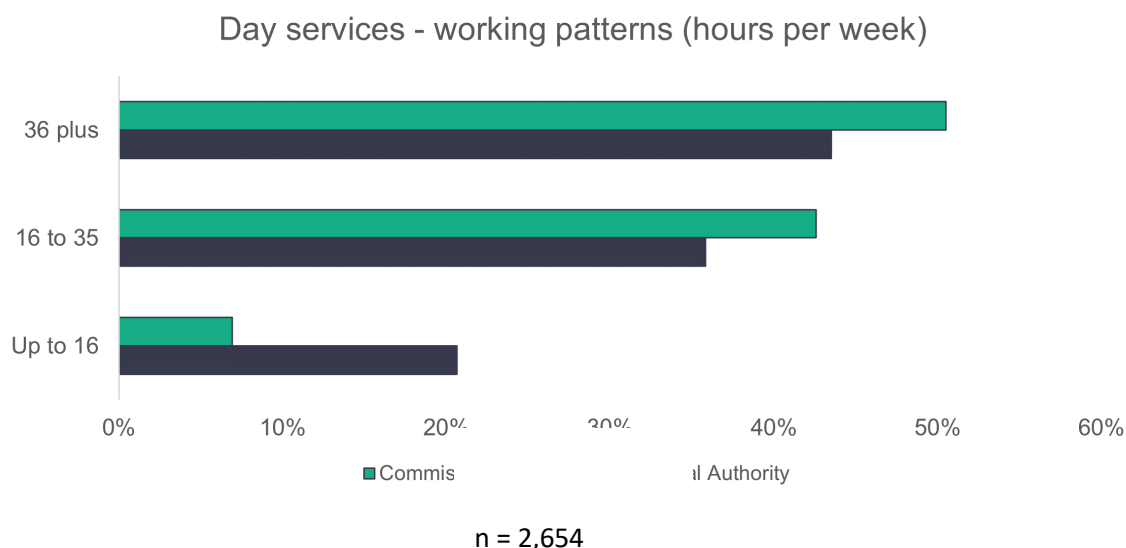
3.4.7 Day services – vacancies

Out of an estimated total day services workforce of 2,954, there were 218 vacancies reported, and 38 of these were held vacant by employers. Last year, the number of vacancies was 280, with 57 held vacant.

3.5 Supported living services

This year, local authority-run services had a 100 per cent return rate, in comparison to last year's 90.6 per cent. Commissioned services also had an increased return rate, up from 76.3 per cent in 2021 to 90.0 per cent this year. The estimated workforce for the supported living sector is 6,893, which is split between 970 for local authorities and 5,923 for commissioned services.

Local Authority

100%

Returns (%) ▪ Not Returned (%)

n = 12

Commissioned Provider
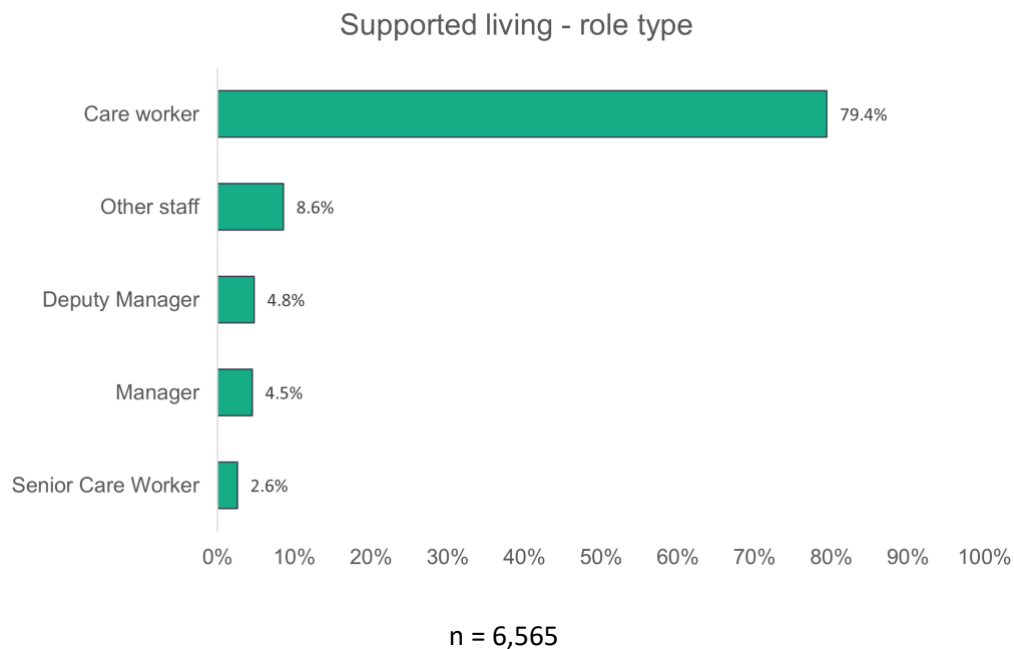
10%

90%

Returns (%) ▪ Not Returned (%)

n = 39
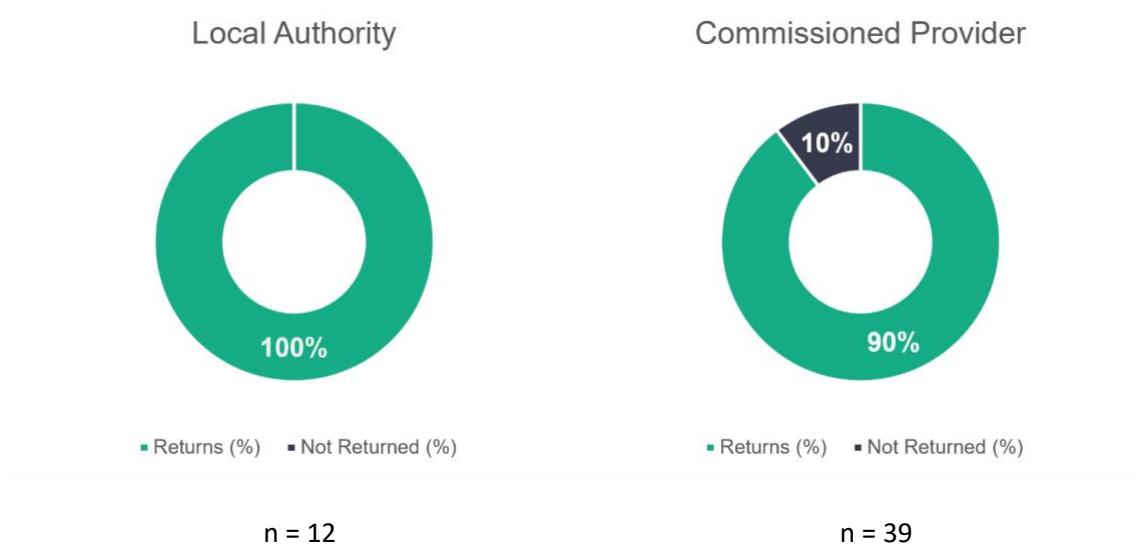
3.5.1 Supported living services – role type

Care workers make up 79.4 per cent of the supported living care workforce, as shown in the following chart.

Supported living - role type

Role	%
Care worker	79.4%
Other staff	8.6%
Deputy Manager	4.8%
Manager	4.5%
Senior Care Worker	2.6%

n = 6,565

3.5.2 Supported living services – gender

Three quarters of those working in supported living are female. This is slightly lower than we'd see generally in social care in Wales. Fewer than five people were reported as gender fluid.
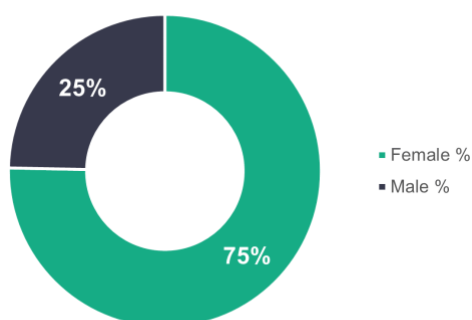
n = 5,972

3.5.3 Supported living services – age profile

Local authorities have an older supported living workforce, with 65 per cent aged 46 or older. Those aged 25 or under were more than twice as likely to work in commissioned services than a local authority supported living scheme.
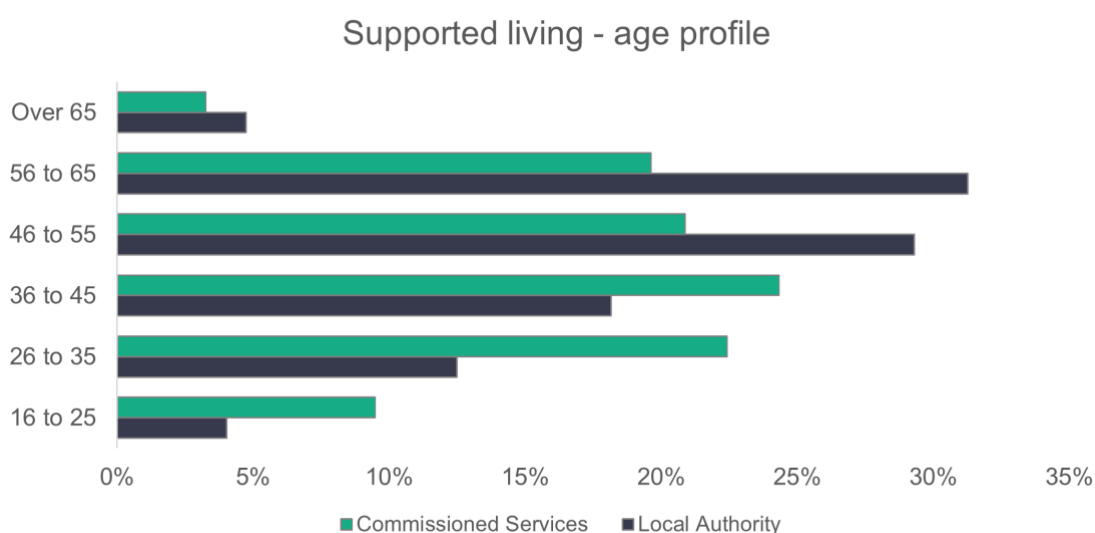
n = 5,933

3.5.4 Supported living services – ethnicity

The ethnicity of those working in supported living services broadly reflects the ethnic diversity of the Welsh population. The marked exception is the larger proportion of black workers in the sector, who make up one in 18 of the supported living workforce.

Supported living - ethnicity
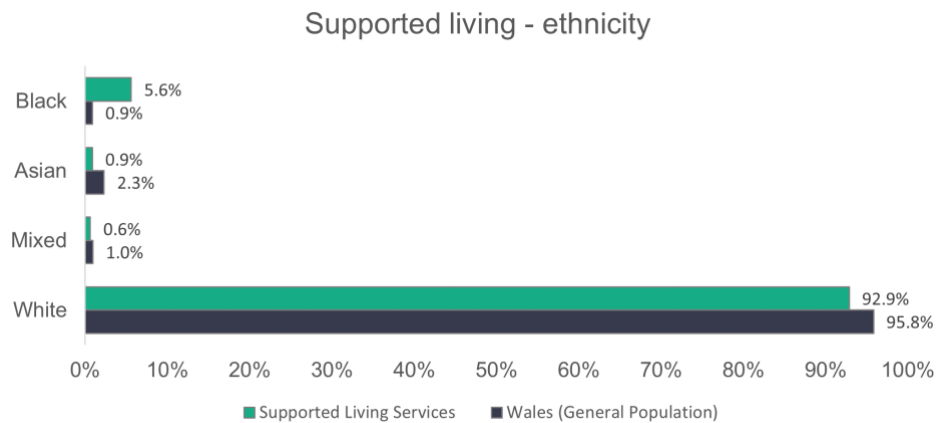
Ethnicity	Supported Living Services	Wales (General Population)
Black	5.6%	0.9%
Asian	0.9%	2.3%
Mixed	0.6%	1.0%
White	92.9%	95.8%

■ Supported Living Services   ■ Wales (General Population)

n = 4,225

3.5.5 Supported living services – contract type

Most supported living service workers are employed on permanent contracts. This is the case for 93.7 per cent of those employed in commissioned services and 70.2 per cent of those employed directly by local authorities. Almost a quarter (24.3 per cent) of local authority workers are employed on zero-hours or casual contracts, compared to a lower percentage (4.4 per cent) of commissioned service workers.

Supported living - contract type

■ Commissioned Services   ■ Local Authority
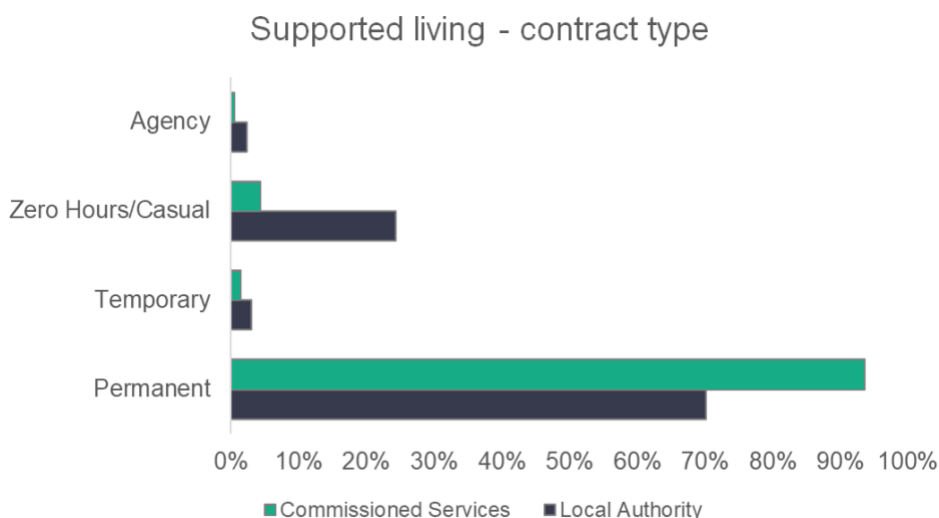
n = 5,997

3.5.6 Supported living services – working patterns

Workers in both local authority and commissioned services span a wide range of working hours, with almost half (48 per cent) of those in the commissioned services sector working full-time hours (the corresponding figure for local authorities is 27.1 per cent). A much bigger proportion of people work fewer than 16 hours in local authorities compared to commissioned services.

Supported living - working patterns (hours per week)
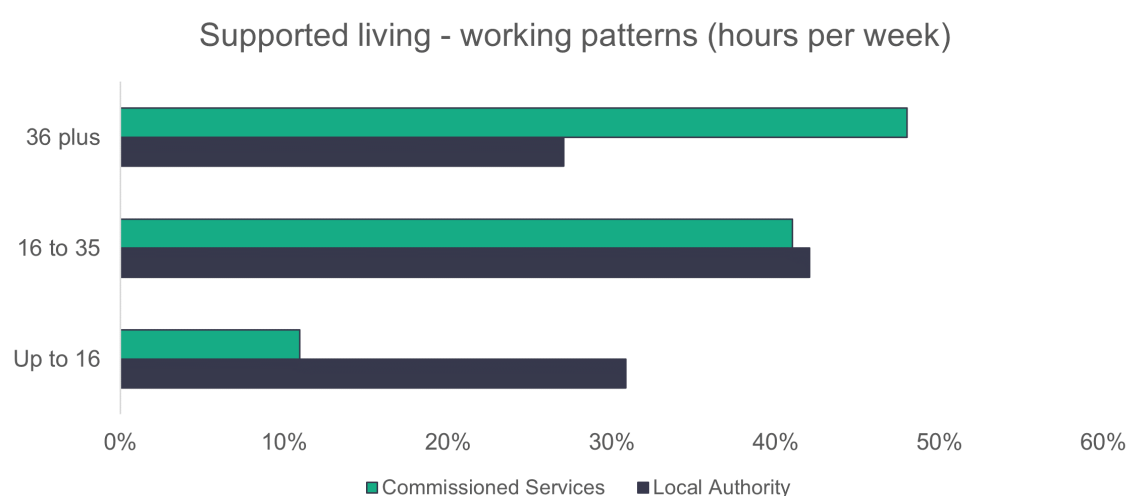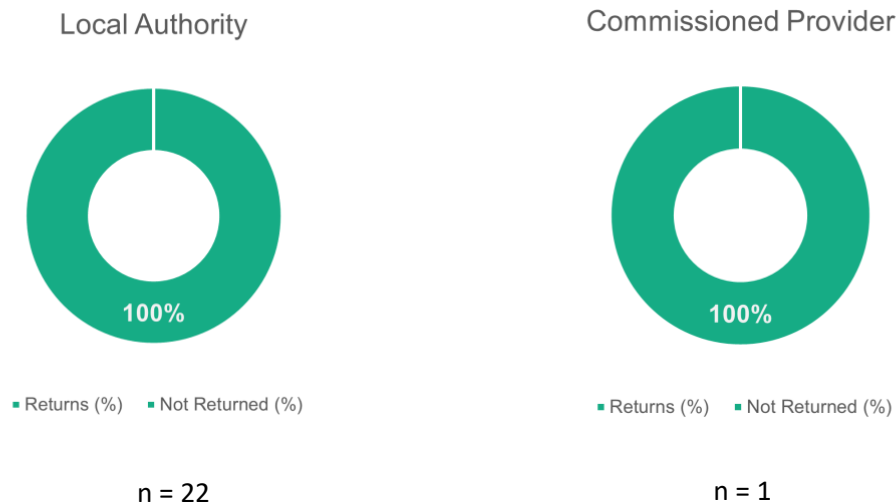

n = 5,884

3.5.7 Supported living – vacancies

The drop in workforce from 7,750 in 2021 to 6,893 this year is reflected in the number of vacancies, with 714 supported living vacancies in 2022 compared to 264 in 2021.

3.6 Social work teams

Return rates for both local authority-run services and commissioned services were 100 per cent, as seen below.

Local Authority

100%

■ Returns (%) ■ Not Returned (%)

n = 22

Commissioned Provider

100%

■ Returns (%) ■ Not Returned (%)

n = 1

3.6.1 Social work teams – role type

The survey shows there are 4,181 registered social workers working in frontline services, accounting for 48.8 per cent of the social work teams' workforce.

Social work teams - role type

Role	%
Registered Social Workers (all)	48.8%
Support Worker	27.9%
Other staff	18.5%
Occupational Therapist	2.8%
Speech and Language Therapist	1.1%
Occupational Therapy Assistant	0.8%
Registered Nurse	0.0%

n = 8,566
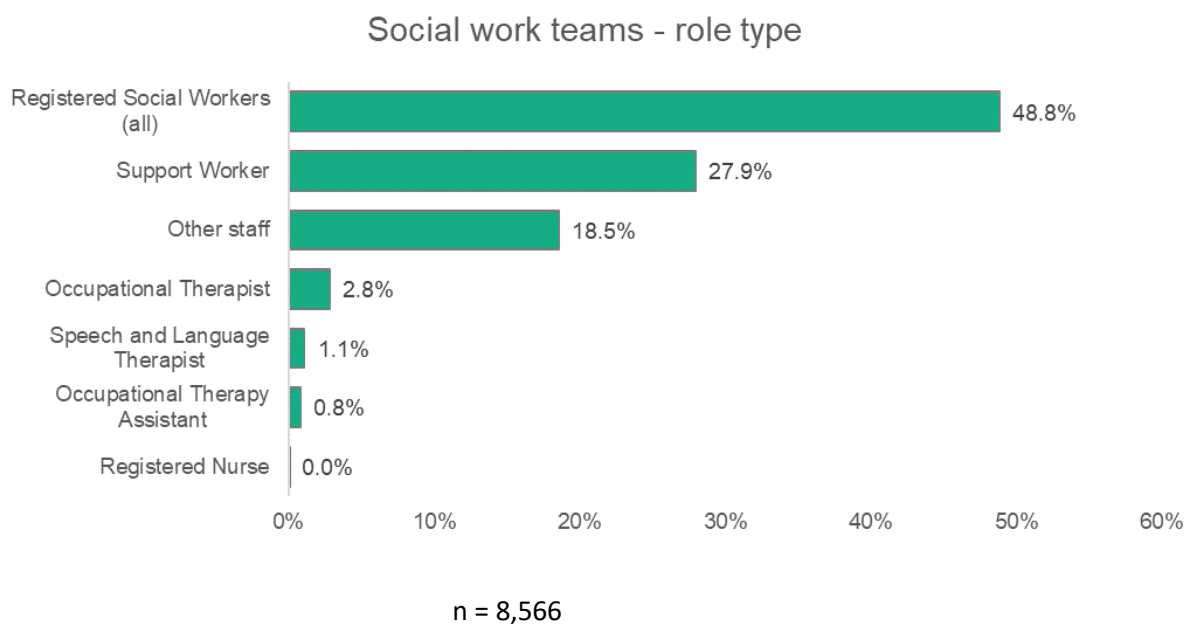
The experience level of the social work teams is shown on the graph below.

Social work teams - level of experience

Role	Percentage
Social Worker - 3 yrs plus	46.4%
Senior Practitioner	16.9%
Team Manager	13.6%
Social Worker - 1-3 yrs	12.8%
Deputy Team Manager	4.8%
Social Worker - FYIP	4.4%
Consultant Social Worker	1.0%

n = 4.181
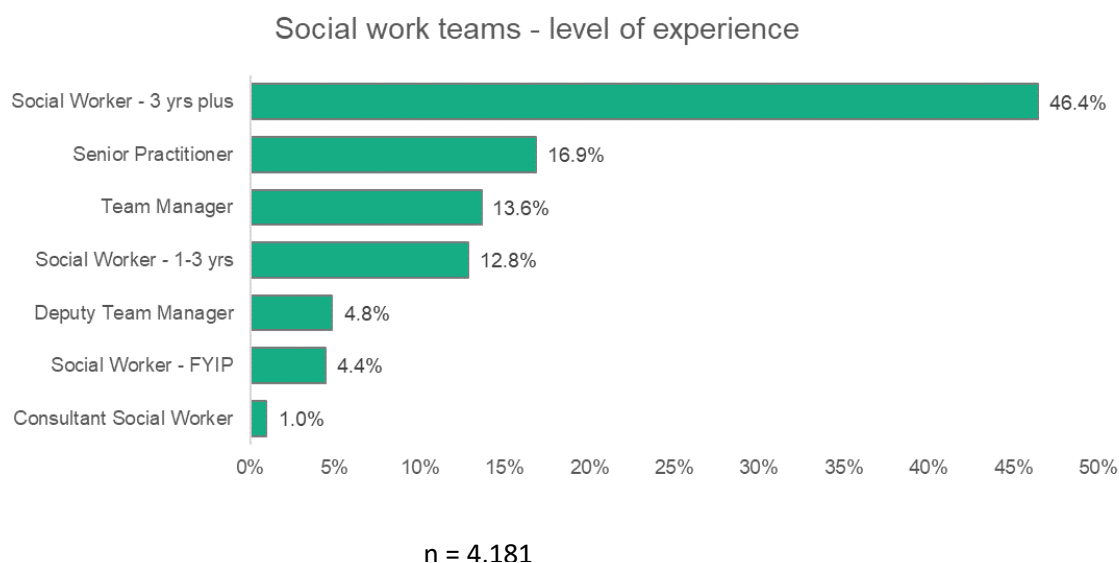
Nearly half (46.4 per cent) of all social workers reported have been qualified for three years or longer, with a further 16.9 per cent in senior practitioner roles and 18.4 per cent in managerial roles. Newly qualified social workers (first year in practice) make up 4.4 per cent of the total registered social worker numbers employed in frontline services.

### 3.6.2 Social work teams – gender

The gender split for social work teams is broadly in line with what we see across the whole social care workforce in Wales, with women occupying most posts at 85 per cent.

- Female %
- Male %

15%

85%
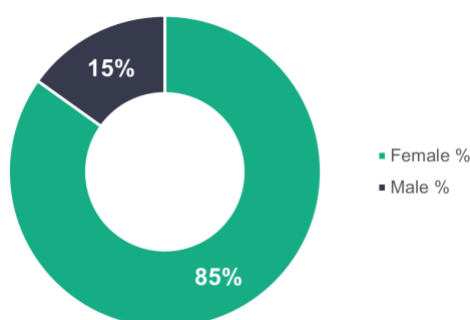
n = 8,529

3.6.3 Social work teams – age profile

Age distributions for workers in frontline services are in line with what we see generally, with local authority staff older than those in commissioned services.

Social work teams - age profile
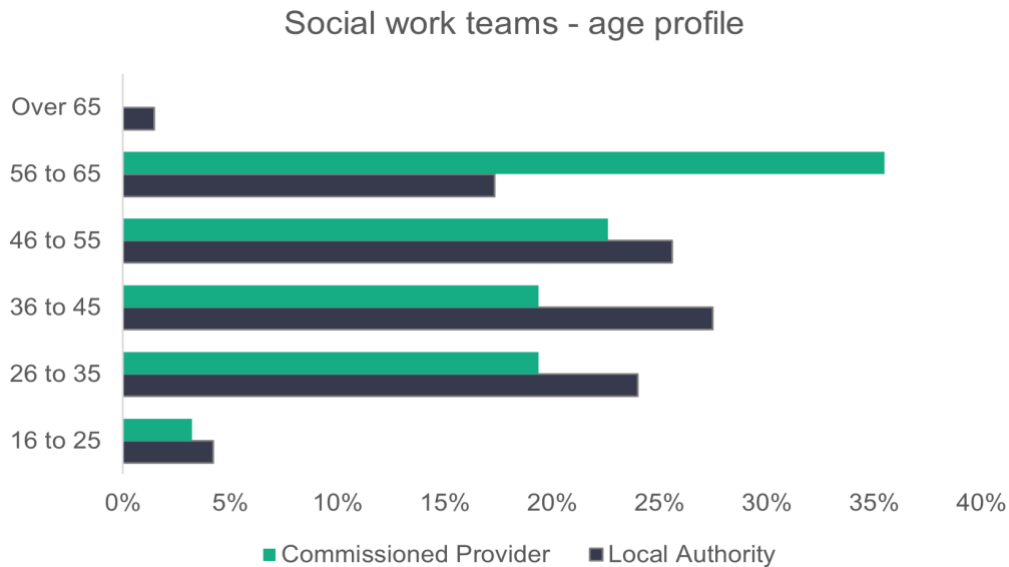
n = 8,549

3.6.4 Social work teams – ethnicity

The frontline social work population appears to be less diverse than the Welsh average. The proportion of white workers is 1.2 per cent higher than the average, while the percentage of Asian workers is 1.7 per cent below the Welsh average.

Social work teams - ethnicity
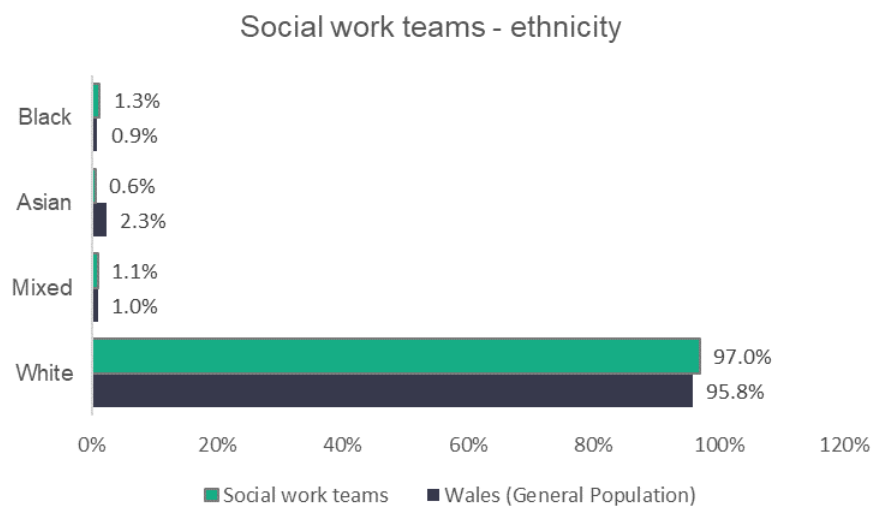
n = 6,519

3.6.5 Social work teams – contract type

All workers from commissioned services and 82.7 per cent of local authority-led services are employed on permanent contracts.

Social work teams - contract type

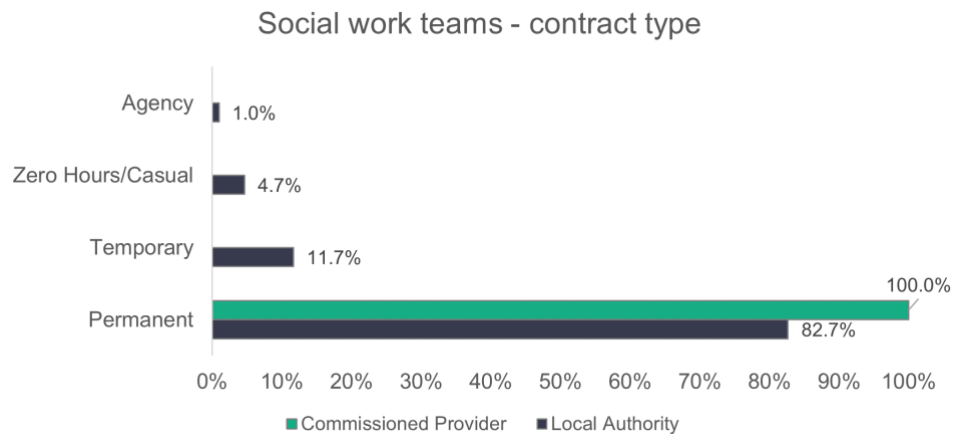| Contract type	Value
Agency	1.0%
Zero Hours/Casual	4.7%
Temporary	11.7%
Permanent (Commissioned Provider)	100.0%
Permanent (Local Authority)	82.7%

■ Commissioned Provider ■ Local Authority

n = 8,566

3.6.6 Social work teams – working patterns

Most frontline workers (72.5 per cent) are contracted as full-time staff in local authority-led services, and 54.8 per cent of frontline workers in commissioned services are employed on full-time contracts.

Social work teams - working patterns (hours per week)

■ Commissioned Provider ■ Local Authority
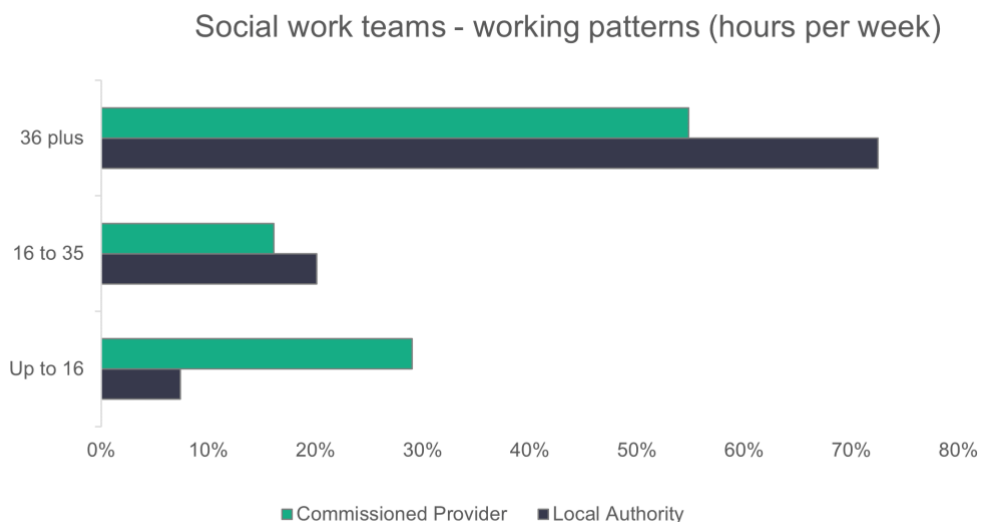
n = 8,428

3.6.7 Social work teams – vacancies and turnover

Frontline teams reported 703 vacancies (451 in children's social work teams, 243 in adult social work teams and five in teams that work with both adults and children). That makes up 16.5 per cent of the social care vacancies in Wales and represents a vacancy rate of 16.8 per cent of the social work team workforce.

We must emphasise that these vacancies may cover any role type in frontline social work teams, and not just qualified social workers.

Of the vacant posts in frontline services in Wales, 239 (34.0 per cent) were held vacant by employers.

Data suggests that 802 people joined social work teams in 2022 and 686 left, giving a net increase of 116 staff.

When we only consider qualified social worker posts, these data show 544 joiners and 468 leavers, giving a net increase of 76.

3.7 Children's residential care

Local authorities had a 100 per cent return rate for both this and last year. Commissioned services had a return rate of 36.9 per cent, which is a significant decrease from 71.3 per cent for last year.
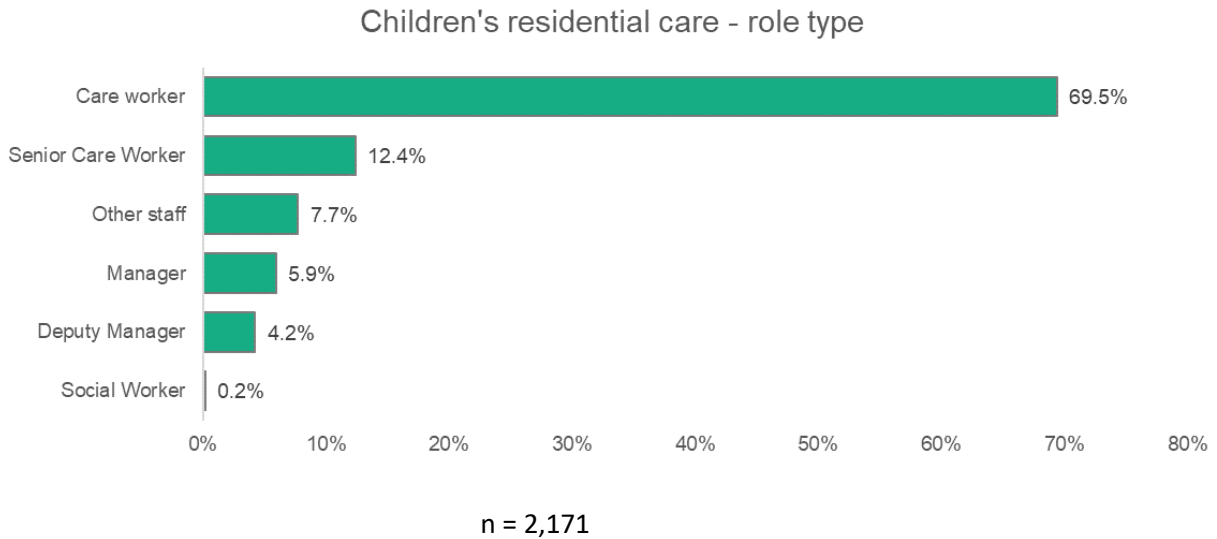
But, the 154 settings for commissioned services represents a 100 per cent increase in the number of settings that reported last year (74). The corresponding number of settings has only slightly changed from 33 last year to 37 this year for local authorities.

In line with the increase in the number of settings for commissioned service providers, the predicted workforce size for commissioned services has also almost trebled from 1,232 last year to 3,595 this year. The combined predicted local authorities and commissioned service providers workforce has increased from 1,838 last year to 4,411 this year.

n = 16   n = 84
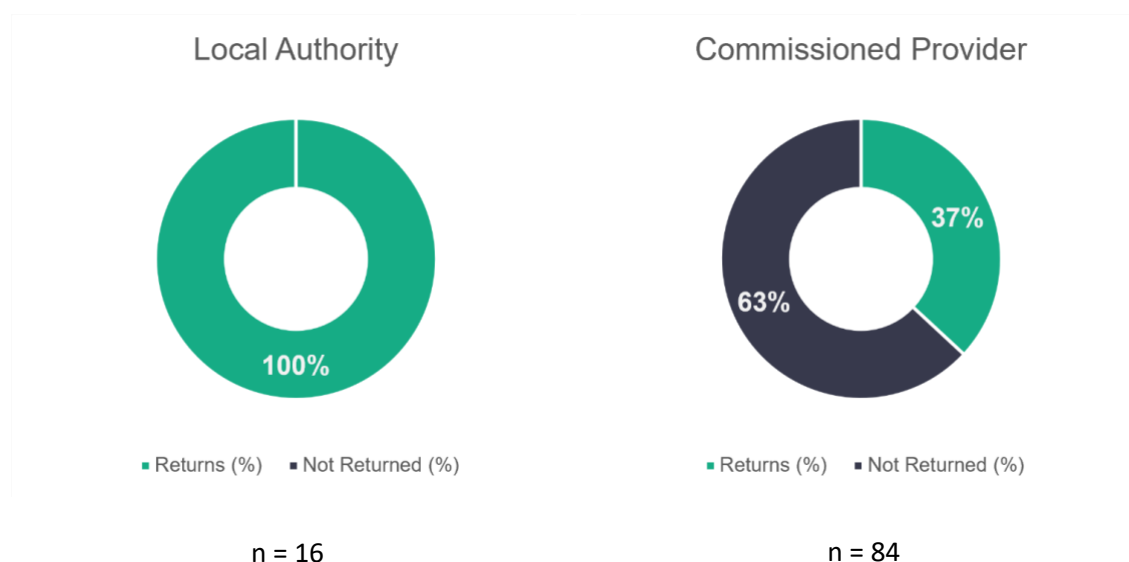
3.7.1 Children's residential care – role type

Care workers make up 69.5 per cent of the children's residential care workforce, as shown below. The next largest category is senior care worker. Other staff include ancillary and business support staff.

Children's residential care - role type

Role	Percentage
Care worker	69.5%
Senior Care Worker	12.4%
Other staff	7.7%
Manager	5.9%
Deputy Manager	4.2%
Social Worker	0.2%

n = 2,171

3.7.2 Children's residential care – gender

More women work in children's residential care, but the proportion of men in the sector is higher than that for the whole social care workforce in Wales. There are almost two men to every three women working in the children's residential care sector. This is significantly different to what we generally see in social care in Wales.

- Female % 62%
- Male % 38%
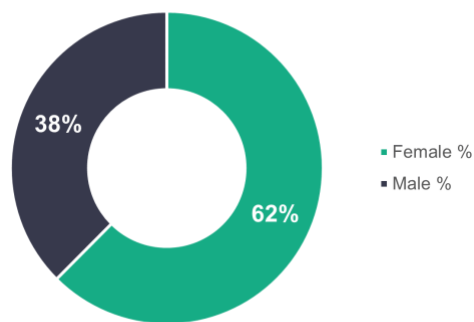
n =1,769

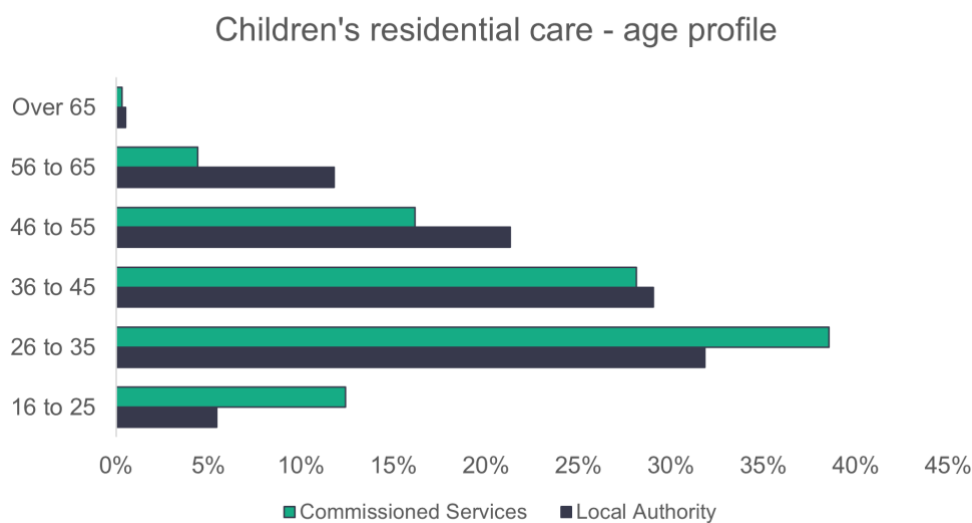### 3.7.3 Children's residential – age profile

The age profile for children's residential care indicates a younger workforce, with the dominant group aged between 26 and 35 in both local authority (31.9 per cent) and commissioned services (38.6 per cent).

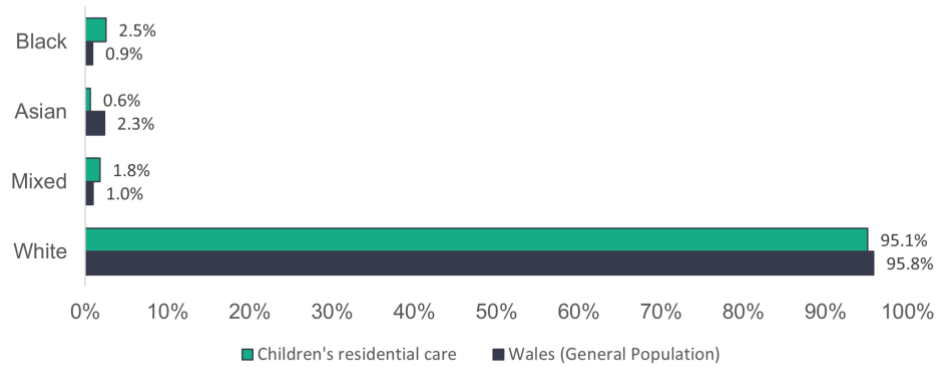**Children's residential care - age profile**

n = 1,740

3.7.4 Children's residential care – ethnicity

There are small differences in the ethnicity of people working in children's residential care in Wales to the general population in Wales. There are fewer Asian people than we'd expect to see, but more black people.

Children's residential care - ethnicity

Black: 2.5% (Children's residential care), 0.9% (Wales General Population)
Asian: 0.6% (Children's residential care), 2.3% (Wales General Population)
Mixed: 1.8% (Children's residential care), 1.0% (Wales General Population)
White: 95.1% (Children's residential care), 95.8% (Wales General Population)

Legend: ■ Children's residential care ■ Wales (General Population)

n = 1,635

3.7.5 Children's residential care – contract type

The commissioned service provider workers in children's residential care are almost all on permanent contracts, making up 95.6 per cent of the workforce. The corresponding proportion of workers employed by local authorities is 69.1 per cent, while 26.3 per cent are on zero-hours or casual contracts.

Children's residential care - contract type

Legend: ■ Commissioned Services ■ Local Authority
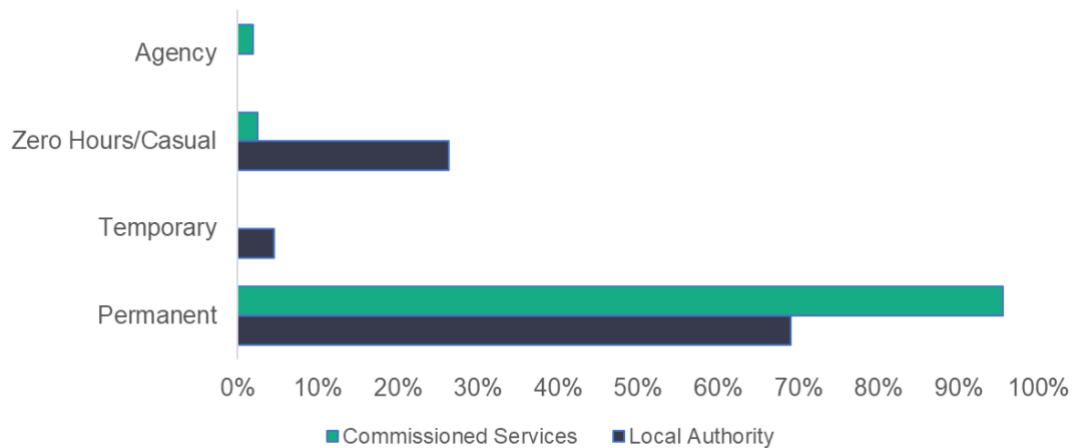
n = 1,763

3.7.6 Children's residential care – working patterns

There are differences between local authorities and commissioned services when it comes to working patterns in children's residential care provision. In commissioned services, 86.1 per cent of the workforce is on full-time hours with only 3.6 per cent working less than 16 hours. The working hours are more spread out for local authority workers, with 38.1 per cent working full-time.

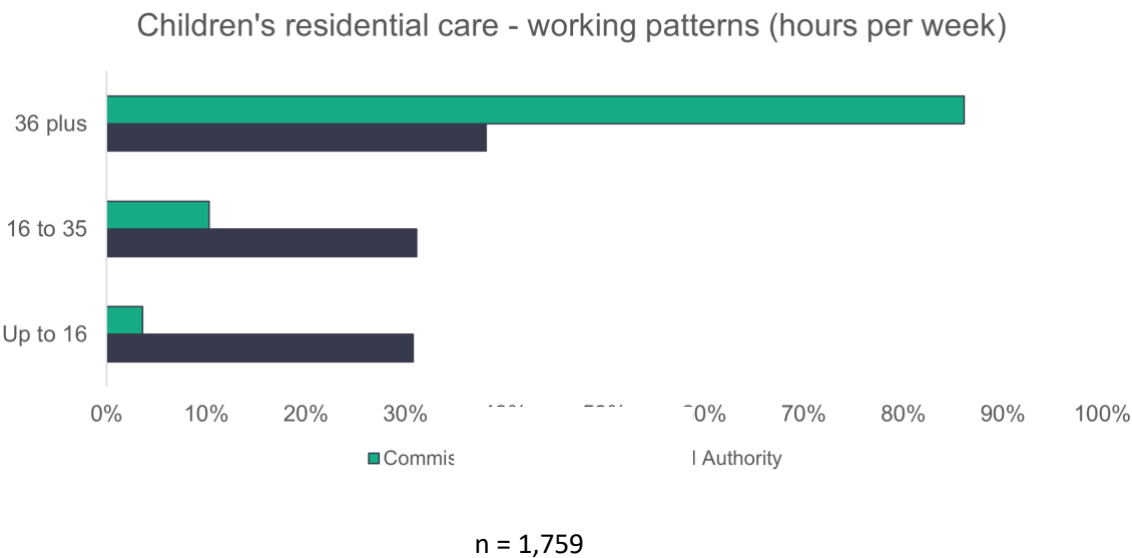Children's residential care - working patterns (hours per week)
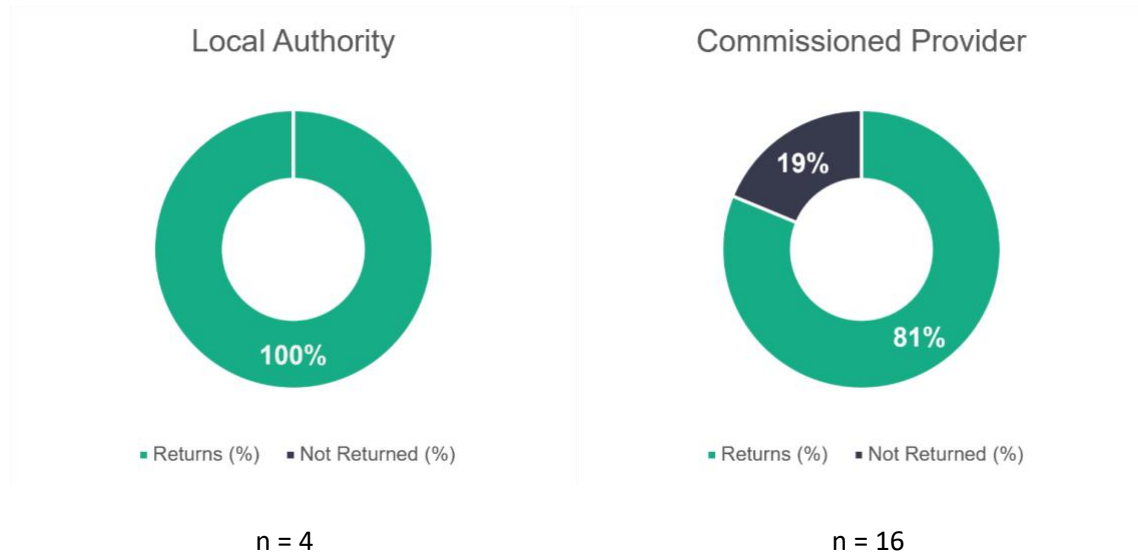

n = 1,759

3.7.7 Children's residential care – vacancies

There were 201 vacancies reported in children's residential care and 67 posts held vacant, representing 3.9 per cent of the overall social care vacancies in Wales. In 2021, the corresponding figures were 141 vacancies with 61 posts held vacant.

3.8 Mental health residential care

Return rates for local authority-run services were higher than those from commissioned services, as we see below.

Local Authority

100%

■ Returns (%) ■ Not Returned (%)

n = 4

Commissioned Provider

19%

81%

■ Returns (%) ■ Not Returned (%)

n = 16

3.8.1 Mental health residential care – role type

Care workers make up 63.1 per cent of the mental health residential care workforce, as shown below. The next largest category is 'other', which is made up of both admin and auxiliary staff.

Residential care, mental health - role type

Role	Percentage
Care worker	63.1%
Other staff	14.6%
Senior Care Worker	10.0%
Manager	5.3%
Deputy Manager	4.0%
Registered nursing staff	2.9%

n = 548
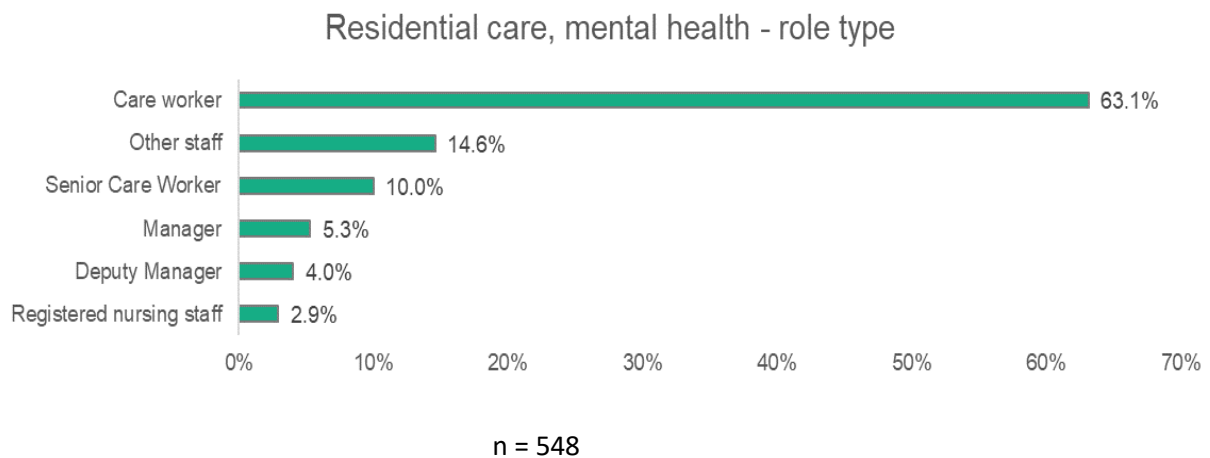
3.8.2 Mental health residential care – gender

The number of women and men working in mental health residential care is broadly in line with what we see across the whole social care workforce in Wales, with women occupying most posts. There's a higher proportion of non-binary staff in mental health residential care, at 0.9 per cent of the workforce.
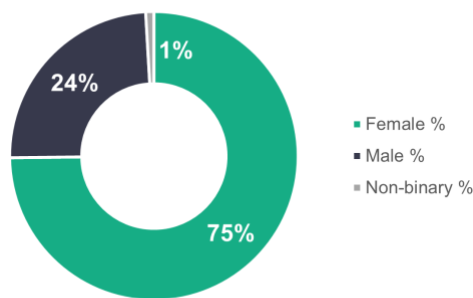
n = 548

3.8.3 Mental health residential – age profile

The proportion of people in older age groups (46 and over) is higher in local authority-run services than in commissioned services. As in other sectors, there's a clear difference in age profile between commissioned and local authority-run services, with younger age groups making up a higher proportion of the workforce in commissioned services.

Residential care, mental health - age profile
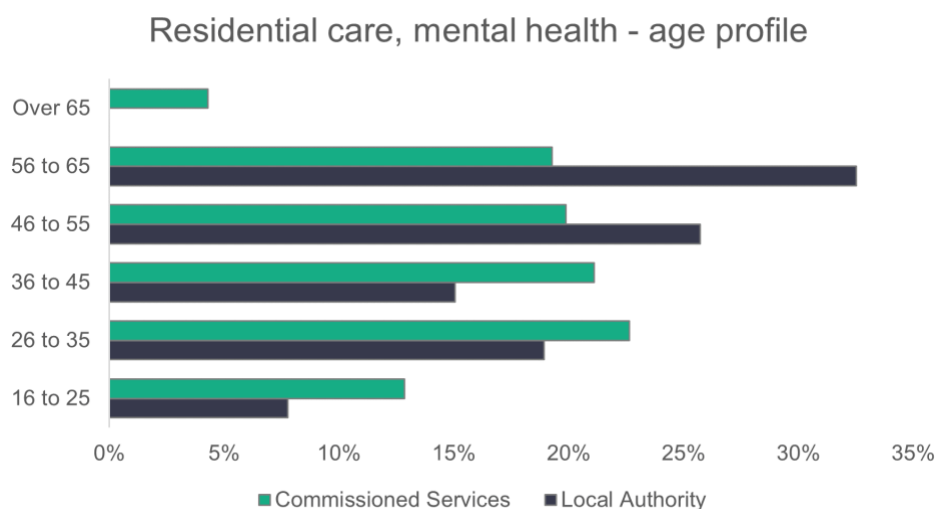
n = 533

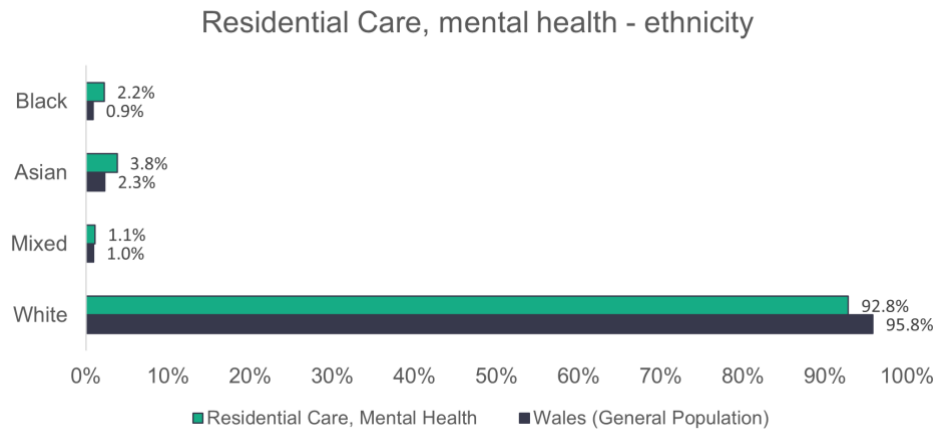## 3.8.4 Mental health residential care – ethnicity

Residential care workers in mental health services are slightly more likely to be from a diverse ethnic background than we'd see in the general population in Wales.

Residential Care, mental health - ethnicity

Ethnicity	Residential Care, Mental Health	Wales (General Population)
Black	2.2%	0.9%
Asian	3.8%	2.3%
Mixed	1.1%	1.0%
White	92.8%	95.8%

n = 447

3.8.5 Mental health residential care – contract type

Residential care, mental health - contract type

Categories: Agency, Zero Hours, Temporary, Permanent
Series: Commissioned Services, Local Authority
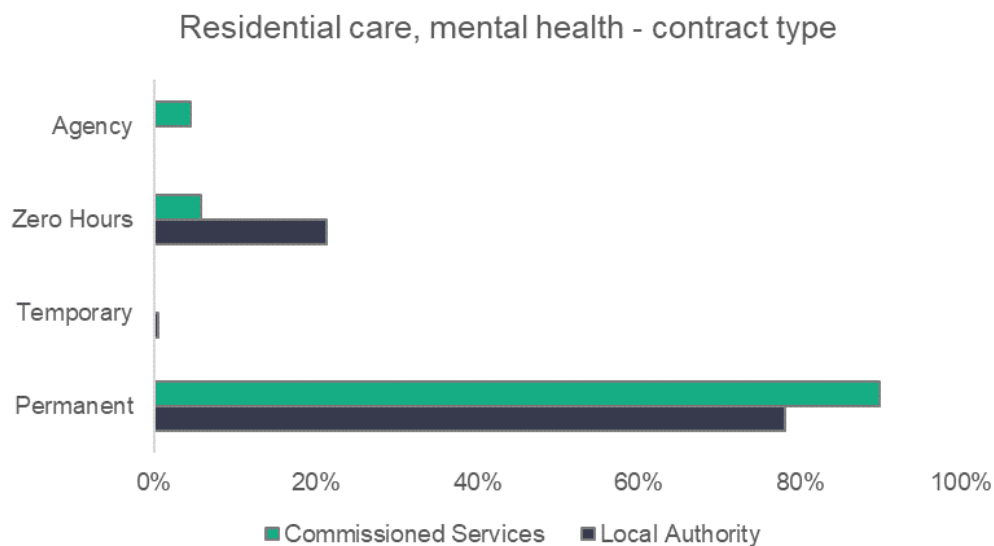
n = 539

In local authority-run services, 78.2 per cent of staff are employed on permanent contracts, compared to 89.8 per cent of those in commissioned services. Over 21

per cent of workers in local authority-run residential mental health services are employed on zero-hours or casual contracts, compared to just over 10 per cent in commissioned services.
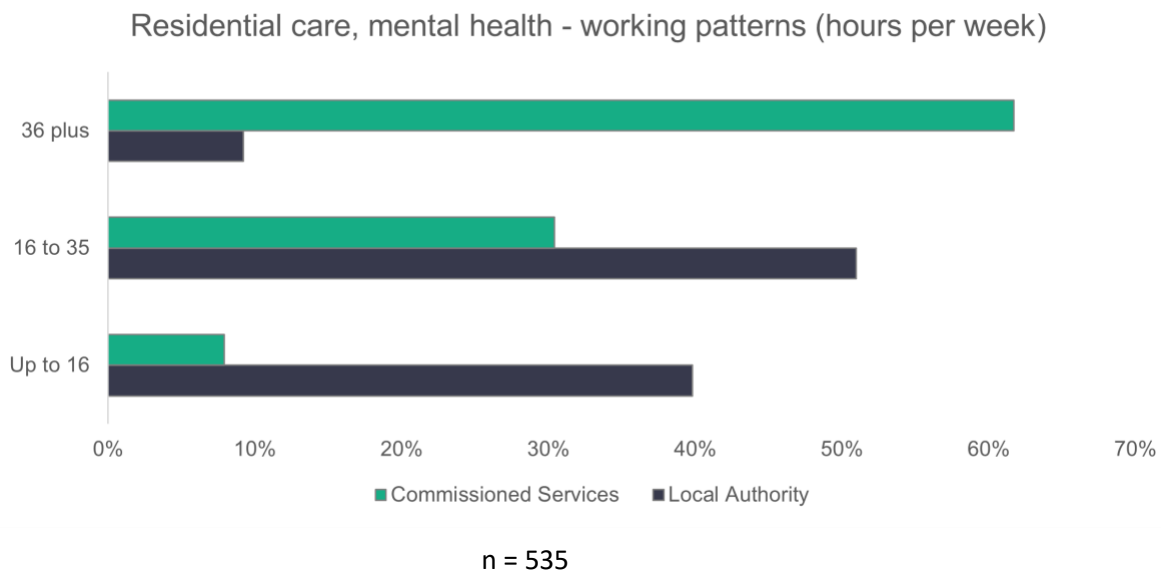
3.8.6 Mental health residential care – working patterns

A much higher proportion of workers in commissioned services (61.7 per cent) are employed full-time. This compares to just 9.2 per cent in local authorities. Two fifths (39.8 per cent) of people in local authority-run services are contracted to work up to 16 hours per week, compared to 7.9 per cent in commissioned services.

Residential care, mental health - working patterns (hours per week)

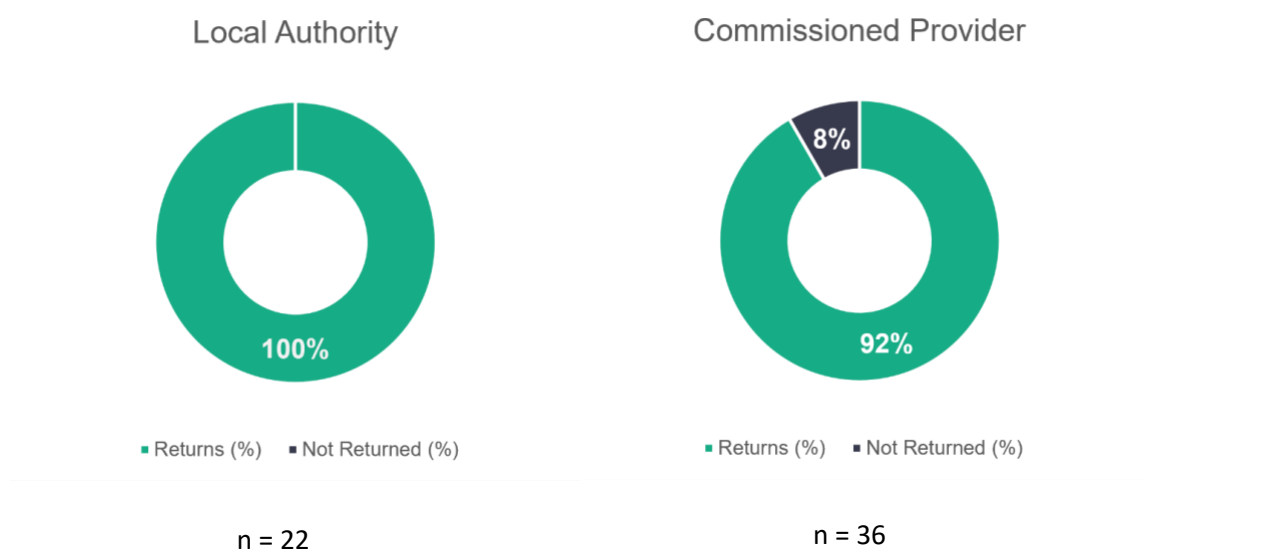■ Commissioned Services ■ Local Authority

n = 535

3.8.7 Mental health residential care – vacancies

The survey showed 64 vacant posts in residential services for mental health in Wales. This represents 1.5 per cent of all social care vacancies in Wales.
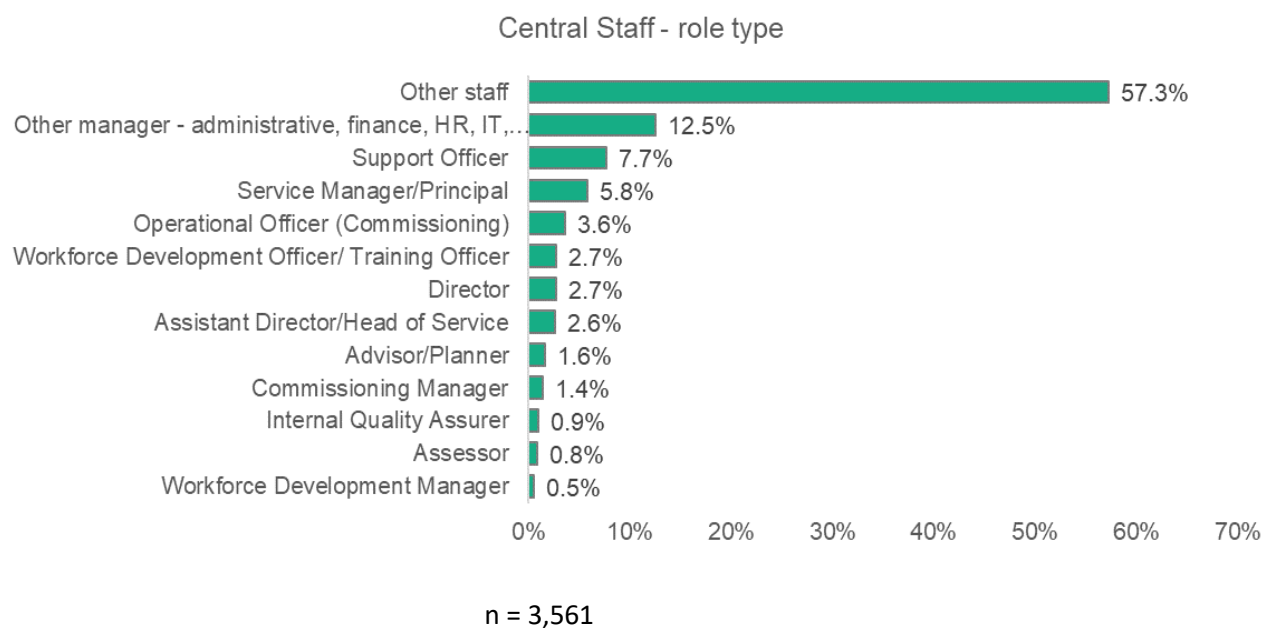
3.9 Central staff

Central services are made up of back office, business support, domestic, catering and other ancillary posts needed by organisations as part of their day-to-day operations.

Local Authority

100%

- Returns (%) - Not Returned (%)

n = 22

Commissioned Provider

8%

92%

- Returns (%) - Not Returned (%)

n = 36

3.9.1 Central staff – role type

Other staff includes business support and auxiliary workers and makes up 57.3 per cent of the central staff workforce, as shown below. We'll work with our data providers to learn more about the role types included under other staff so we can improve the way we group the data.

Central Staff - role type

Role	%
Other staff	57.3%
Other manager - administrative, finance, HR, IT,...	12.5%
Support Officer	7.7%
Service Manager/Principal	5.8%
Operational Officer (Commissioning)	3.6%
Workforce Development Officer/ Training Officer	2.7%
Director	2.7%
Assistant Director/Head of Service	2.6%
Advisor/Planner	1.6%
Commissioning Manager	1.4%
Internal Quality Assurer	0.9%
Assessor	0.8%
Workforce Development Manager	0.5%

n = 3,561

3.9.2 Central staff – gender

The gender split for central staff shows women occupying most posts, but more men work in central services than we'd expect to see generally in social care in Wales.

Female %
Male %
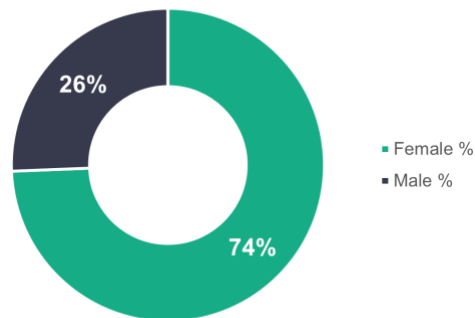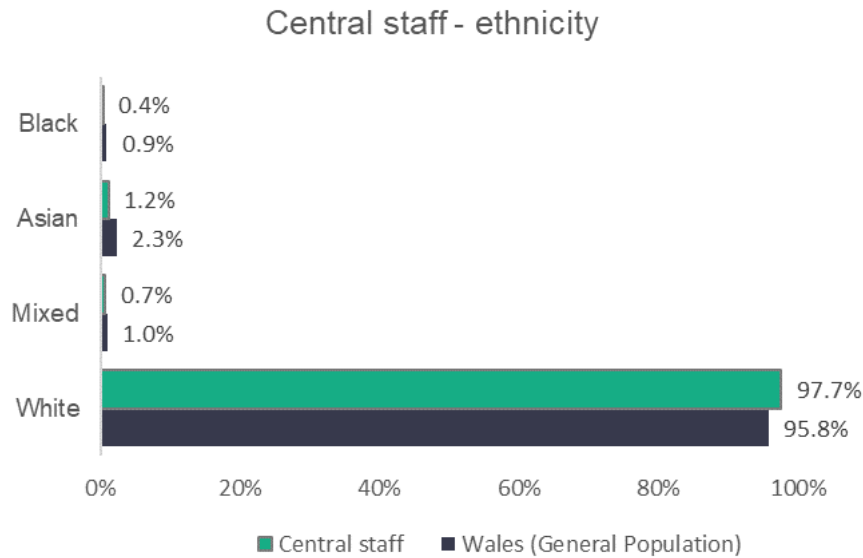
n = 3,438

3.9.3 Central staff – age profile

The age of workers is broadly similar for local authorities and commissioned services. Commissioned service providers have a much higher rate of people over the age of 65.

Central staff - age profile

Commissioned Services   Local Authority
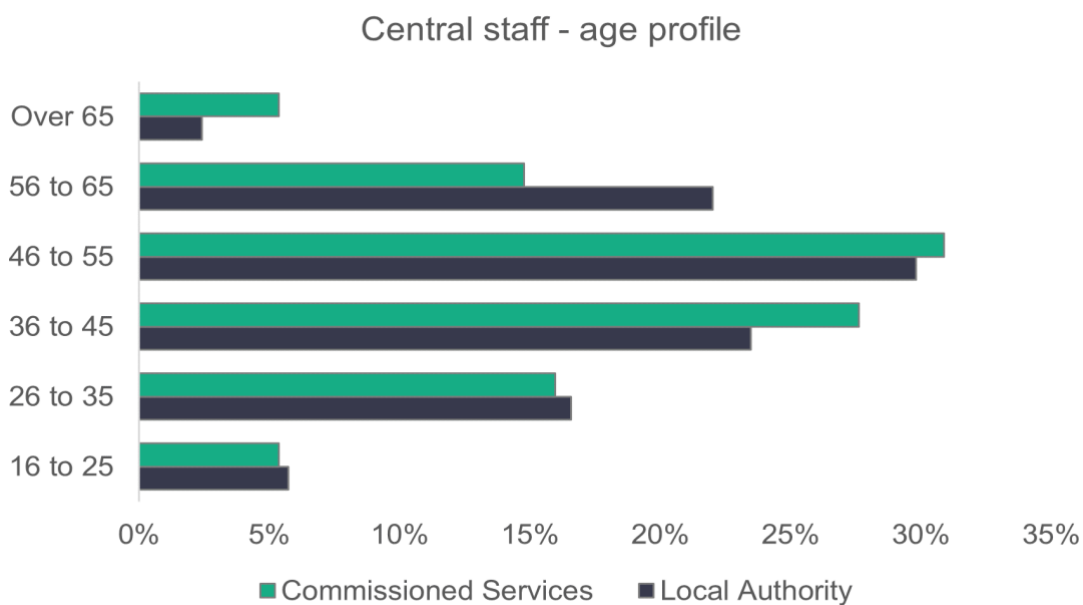
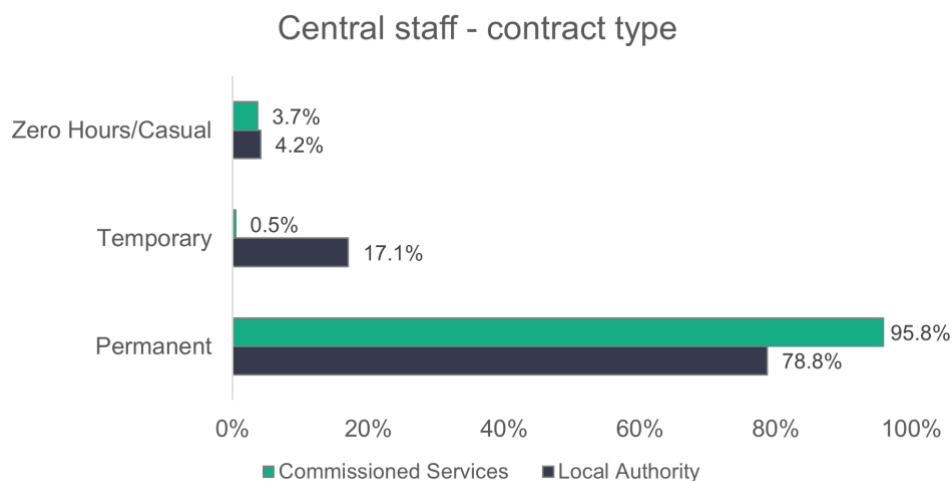n = 3,572

3.9.4 Central staff – ethnicity

The ethnic background of workers in central services mirrors the general population of Wales, with 97.7 per cent of workers reported as being from white backgrounds.

Central staff - ethnicity

Ethnicity	Central staff	Wales (General Population)
Black	0.4%	0.9%
Asian	1.2%	2.3%
Mixed	0.7%	1.0%
White	97.7%	95.8%

■ Central staff   ■ Wales (General Population)

n = 2,568

3.9.5 Central staff – contract type

We found that 78.8 per cent of employees from local authority-led services are on permanent contracts, compared to 95.8 per cent of people who work for commissioned services.

Central staff - contract type

Contract type	Commissioned Services	Local Authority
Zero Hours/Casual	3.7%	4.2%
Temporary	0.5%	17.1%
Permanent	95.8%	78.8%

■ Commissioned Services   ■ Local Authority

n = 3,564

3.9.6 Central staff – working patterns

Over 70 per cent of employees who work for local authority-led services and commissioned services are employed full-time.

Central staff - working patterns (hours per week)
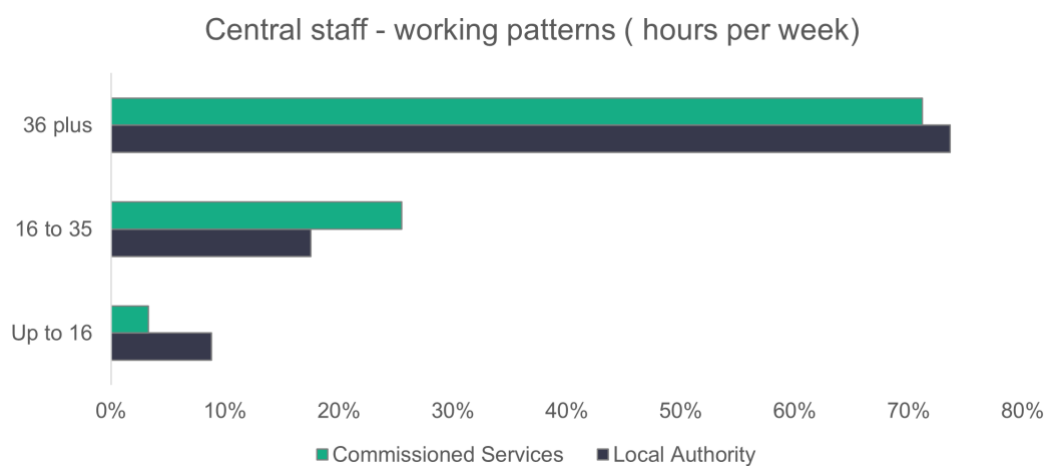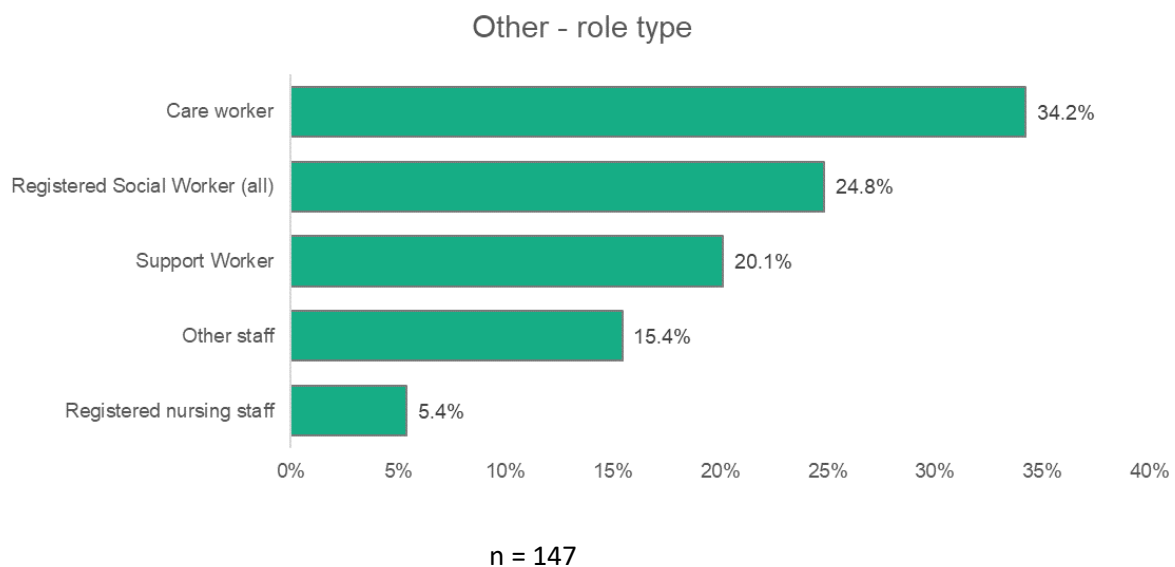
n = 3,584

3.9.7 Central staff – vacancies

Central services have an estimated 141 vacancies, representing 3.3 per cent of all social care vacancies in Wales. This is a 0.9 per cent decrease from 2021, when there were 235 estimated vacancies, which represented 4.2 per cent of all social care vacancies in Wales.

3.10 Other services

For services that don't naturally sit within the listed setting types, we created an 'other' category to make sure we're able to collect the details of everyone who works in social care in Wales.

In 2022, we received information from 10 social care service organisations (nine commissioned services and one local authority), who reported a total of 147 workers, whose roles are shown below.

Other - role type

Role	Percentage
Care worker	34.2%
Registered Social Worker (all)	24.8%
Support Worker	20.1%
Other staff	15.4%
Registered nursing staff	5.4%

n = 147

Although these services are included in our overall estimates of the workforce, we didn't receive enough data to calculate reliable estimates for other attributes of these services. The service type providers most associate with is something we've identified as a possible area for improvement and we'll continue to work with providers to make sure that categorisations used in the data collection accurately reflect the full range of services and professions we have in Wales.

Appendix A - Data categories

Setting type:

- Adult placement schemes
- Residential care – mental health services
- Domiciliary care
- Day services
- Social work teams – children
- Personal assistants
- Central (support) staff
- Residential care – adult
- Residential care – children
- Supported living services
- Social work teams – adults
- Social work teams – all
- Foster carers

Role types:

- Adult placement manager
- Adviser/planner
- Care worker
- Deputy manager
- Director
- Head of service (assistant director)
- Manager
- Occupational therapist
- Other staff
- Personal assistant
- Planner
- Registered nursing staff
- Senior practitioner
- Social care planning and commissioning manager
- Social care planning and commissioning support officer
- Social worker (3+ years)
- Speech and language therapist
- Team manager (social work)
- Workforce development officer/Training officer
- Adult placement worker
- Assessor
- Consultant social worker
- Deputy team manager (social work)
- Foster carer
- Internal quality assurer (IQA)
- Nursing assistant/Auxiliary nurse
- Occupational therapy assistant
- Other manager
- Physiotherapist
- Registered nurse
- Senior care worker
- Service manager/Principal officer
- Social care planning and commissioning operational officer
- Social worker (2 to 3 years)
- Social worker (first year in practice)
- Support worker
- Workforce development manager

Age
- 16 to 24
- 25 to 34
- 35 to 44
- 45 to 54
- 55 to 64
- 65+

Gender
- Male
- Female
- Gender fluid
- Non-binary

Disability
- Yes
- No

Working hours per week

- Up to 16 hours
- 16 to 36 hours
- 36 plus hours

Contract status

- Permanent
- Zero hours / Casual
- Voluntary
- Bank/Agency

Vacancies

- To be filled
- Held vacant

Welsh language proficiency level

- No ability
- Welsh entry
- Welsh foundation
- Welsh intermediate
- Fluent

First language

- Welsh first language
- English first language
- Other first language

Ethnicity

- White Welsh
- White British
- White Irish
- White European
- White and Asian
- White / Black African
- White / Black Caribbean
- White (other)
- Black Welsh
- Black British
- Black Caribbean
- Black African
- Black (other)
- Asian Welsh
- Asian British
- Asian Pakistani
- Asian Indian
- Asian Chinese
- Asian Bangladeshi
- Asian (other)
- Arabic
- Mixed ethnicity
- Mixed (other)
- Gypsy / Traveller
- Other ethnicity
- Prefer not to say

Appendix B - The way we dealt with missing data and estimating the total size of the social care workforce in Wales

One of our main aims is to accurately estimate the size of the social care workforce in Wales. But the dataset created for this study contained both fully and partially completed sections. This crucially included missing entries for the number of people in post, which were needed to calculate the total workforce in Wales. In analysing these data, we developed a way of accounting for missing data to allow us to create estimates for the numbers working in social care in Wales.

This problem is ideally suited for a machine-learning approach to predict the missing inputs for the social care workforce dataset. This type of machine-learning problem is classed as supervised learning[1], where a collection of known values (number of roles), known as labelled values, is given. The goal is to use the labelled data to train a machine-learning model to make accurate predictions for unlabelled data.

We tried several regression algorithms that were suitable for our data. The one that gave us the best results was eXtreme Gradient Boosting (XGBoost). XGBoost is a supervised learning algorithm that can be used for both regression and classification problems and is an implementation of the gradient boosted decision tree algorithm. Boosting is when many weaker learners (decision trees in this case) are combined to create a strong learning model[2]. The term "gradient" in XGBoost refers to the algorithm's use of a parameter optimisation method called gradient descent that aims to minimise errors by adjusting the settings' values, improving the model's performance.

The metrics we chose to assess the model's performance on are the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). RMSE and MAE are both used to measure the distance between the predicted and actual values of the target variable. RMSE emphasises the larger errors and is defined as the square root of the square of the average distance between the predicted and observed values. This means it's higher when the difference between the predicted and observed values is big. MAE only measures the average distance between the predicted and observed values. Even though it doesn't emphasise large errors, it still gives an indication of how different the predictions are on average.

Number-in-post predictions

Around 43.5 per cent of the values relating to the number of people in post were missing from the data and, before machine-learning predictions could be used, we needed to carry out some data cleansing and engineering. We started by

---

[1] The other types of machine learning are unsupervised learning, semi-supervised learning and reinforcement learning and they are explained here: https://link.springer.com/article/10.1007/s42979-021-00592-x.

[2] 2 For more information about XGBoost and decision trees, visit https://www.nvidia.com/en-us/glossary/data-science/xgboost/.

removing the rows where the total number of settings and number of settings reported are equal to zero, as well as the ones for foster care and personal assistants' settings, as this data was of much poorer quality.

We then used the Care Inspectorate Wales dataset to get the total number of settings for some of the organisations that didn't give us this information. As some organisations have the same setting type in different counties, we calculated the total number of people in post for each organisation per authority and setting type and used it to predict the missing values for each organisation that didn't return data for each setting type and its corresponding authority.

We also calculated Pearson's correlation between our continuous variable, total number of settings, and the target variable, number of people in post. From this, we created boxplots to assess whether there might be a relationship between our categorical variables, authority and setting type, and the target variable (number in post). The correlation between our target variable and the total number of settings is 0.5. The correlation coefficient ranges between minus one and one. The closer it is to each end point, the stronger the correlation, while the closer it is to zero, the weaker. A value of 0.5 means there's a moderate positive correlation between the total number of settings and the number in post, so the higher the total number of settings is, the higher the number in post, and vice versa.
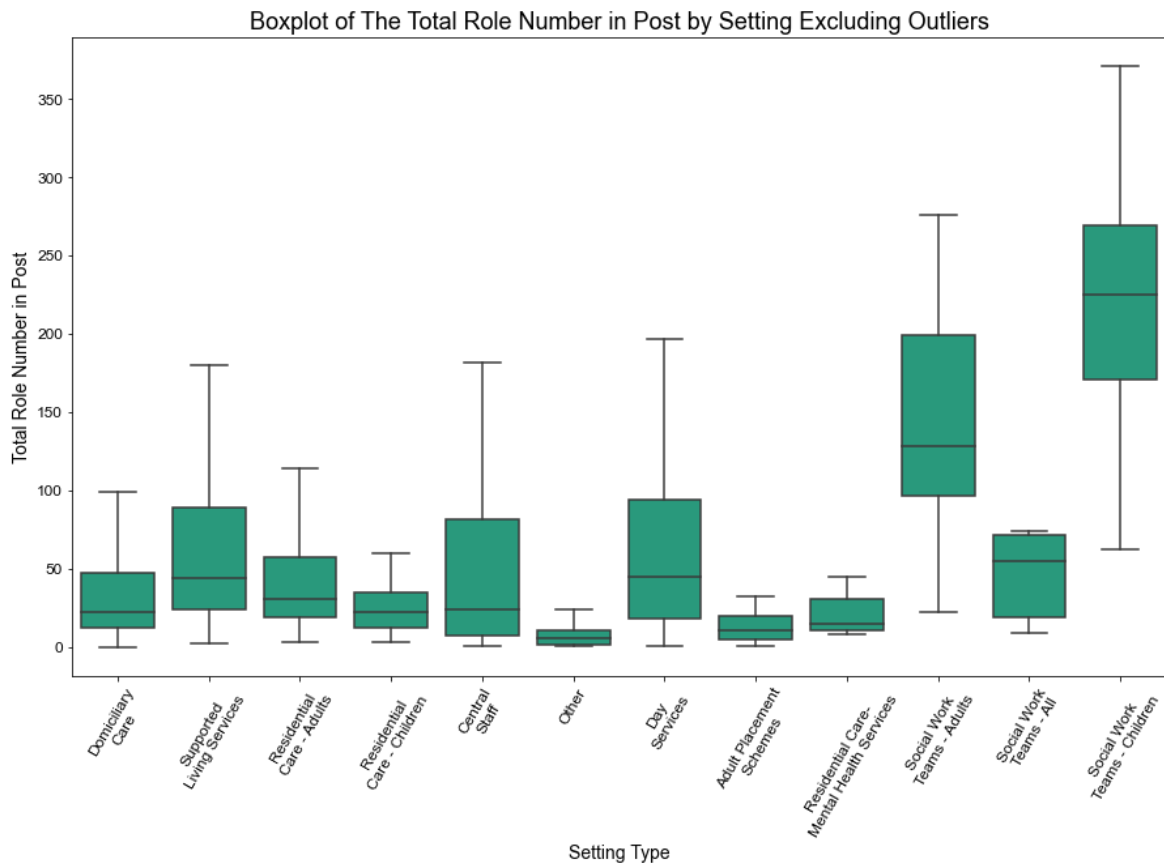
In determining the relationship between our categorical variable, authority, and our target variable, the box plot on the next page suggests that there might be a relationship between the two variables, as the median (the horizontal line in the boxes) of the number in post in some of the authorities is somewhat variable. Outliers were removed before plotting the data.

The box in a boxplot represents the range between the 25th and 75th percentiles of the data and is called the interquartile range. A relationship between the two variables is further suggested when we look at the spread of the total-number-in-post variable in each authority, which we can tell from the size of the box and the length of the whiskers (the lines extending from the top of the box to the highest value and the lowest value, excluding outliers).

We can see that boxes for different authorities have different heights, and the whiskers have different lengths, which indicates that the spread of the total number of people in post in each authority is different. This suggests that there's a difference in the distribution of the total-number-in-post variable between the different authorities, which supports our deduction that there may be a relationship between this categorical variable and the target variable.

Boxplot of The Total Role Number in Post by Authority Excluding Outliers
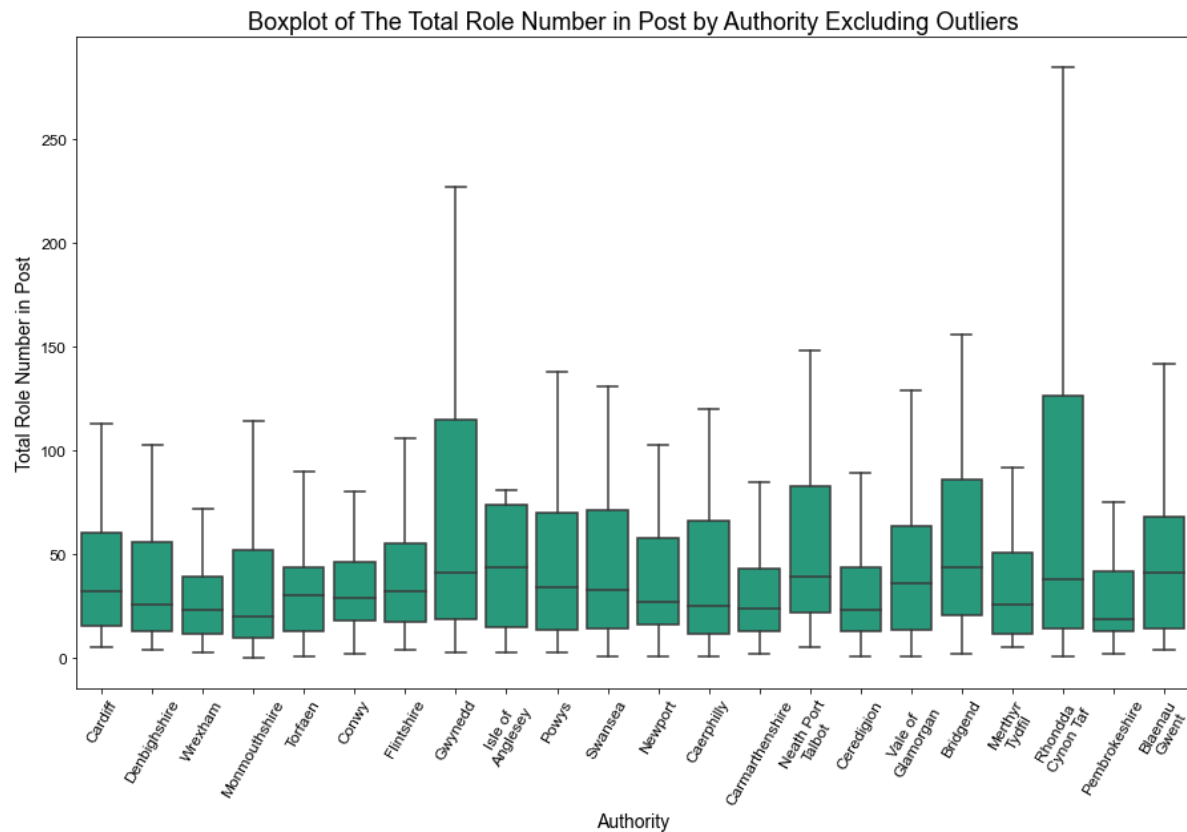
The other relationship we investigated was between our target variable, number in post, and the setting type. It's clear from the box plot below (outliers not shown) that the median of the number in post for each setting is significantly different and is clearer than the previous plot of the authority. This suggests there's a relationship between the two variables. The different lengths of the whiskers and the different heights of the boxes between the settings further suggest there might be a relationship between the two variables.

Boxplot of The Total Role Number in Post by Setting Excluding Outliers

After checking the relationship between our input data and our target variable, we split our data into one dataset that will be used to train and test the model and another that consists of the rows we need to predict the values for. The variables we used as input are authority, setting type and total number of settings.

We used one-hot encoding to convert the nominal categorical variables to a numeric form that can be used in the machine-learning procedure, and we standardised our continuous variables, which improved the model's performance.

Standardisation is transforming numerical data so it has a mean of zero and a standard deviation of one, so our variables have a similar range. It's done by subtracting the mean of the variable from each individual observation and then dividing by the standard deviation. We then split our data into 80 per cent for training and validating the model, and 20 per cent that will be set aside to test its performance on unseen data.

We used grid search to find the best combination of parameters[3] for our XGBoost model and repeated k-fold cross-validation to evaluate its performance. Repeated k-fold cross-validation is a technique of cross-validation where the dataset is divided into k number of subsets of the same size, of which k-1 are used to train the model and the remaining one is used for evaluation. This is repeated multiple

---

3 For more information about grid search please visit https://scikit-learn.org/stable/modules/grid_search.html.

times on randomly selected subsets so that each one is used once for validation, which means that the model evaluation will not be biased.

After tuning our model, it gave us an average MAE value of 0.5 on the repeated k-fold cross-validation and, when tried on the test set, gave us an MAE value of 0.5 and an RMSE of 0.9.

The RMSE value means that the predicted values deviate from the observed ones by 0.9 unit within the range of our standardised target variable.

The MAE value of 0.5 means that the average absolute deviation of the predicted from the observed values is 0.5. We should note that generally, the lower the RMSE and MAE values are, the better the model's performance is.

These values are relatively small compared to our standardised target data range, which indicates that our model's predictions are fairly accurate and that the error level is low to moderate.

## Predicted estimates for vacancies to be filled

For this variable, we only predicted values for the organisations that provided us with data. As in the previous section, we removed foster care and personal assistants' settings from the dataset and then calculated the sum of both the number of people in post and vacancies to be filled for each organisation by authority and setting type.

We then created a Pearson's correlation matrix to see if there are any relationships between our independent continuous variables and our target variable, which is vacancies to be filled.

We found a moderate positive correlation of 0.67 between the sum of vacancies to be filled and the sum of the number in post that we calculated previously. There's a weaker but still moderate positive correlation of 0.51 between the sum of vacancies to be filled and the total number of settings variable.

This positive correlation means that as the total number of settings or the sum of the number of people in post increases, the sum of the vacancies to be filled increases too and vice versa.

When it came to assessing the relationship between the categorical variables and our target variable, we created box plots. We can see from the box plot below of the total role number of vacancies across different authorities that the median values of the total number of vacancies across different authorities are very similar in most cases (Caerphilly, Pembrokeshire and Blaenau Gwent, for example) but are significantly different in some others (Swansea, Isle of Anglesey, for example). This indicates that there may not be a significant relationship between our two variables. But the differences in the whiskers' lengths and the boxes' heights indicate there's a difference in the distribution of vacancies to be filled across the authorities, which suggests that there may be a relationship between our variables.

Boxplot of The Total Number of Vacancies to be Filled by Authority Excluding Outliers
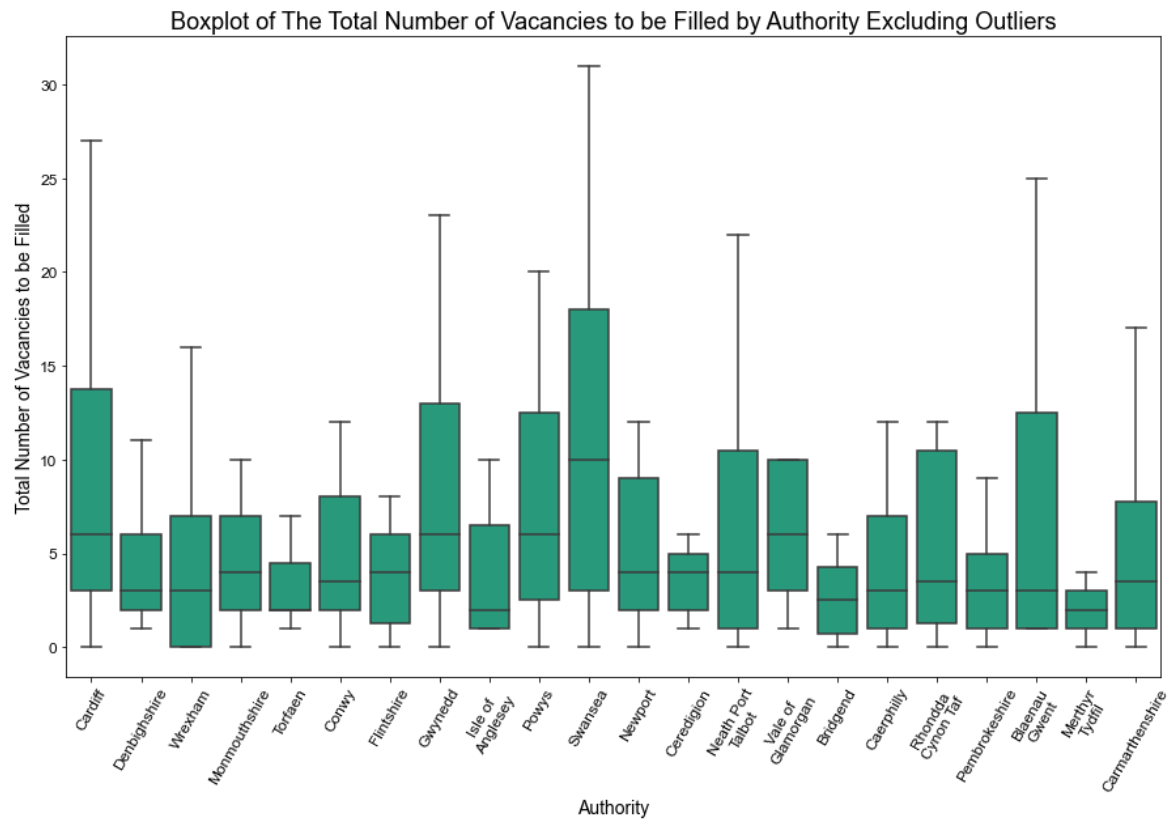
Next, we assessed the relationship between our target variable, number of vacancies to be filled, and the setting type. We can see from the box plot below that the median values across the setting types are very similar, which suggests there might not be a relationship between our target variable and the setting type. But, as we noticed for the authority variable, the lengths of the whiskers and the heights of the boxes are different for the different setting types, which indicates there might be a relationship between the variables.

Boxplot of The Total Number of Vacancies to be Filled by Setting Excluding Outliers
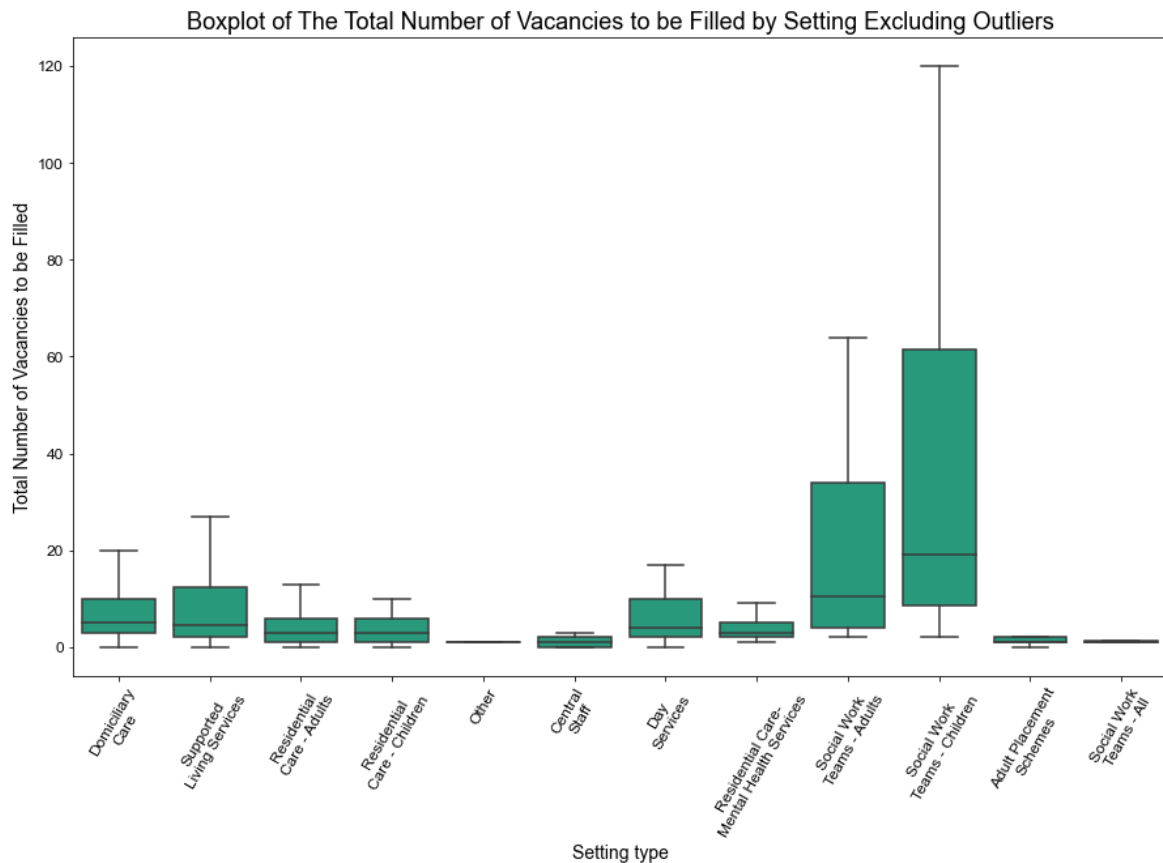
We split our data into a training set and a prediction set, like we did before, with a further split of the training set into 80 per cent for training and 20 per cent for testing. We then used cross-validation with MAE as a metric to evaluate the model's performance on our training data. After knowing the initial average MAE of the repeated k-fold cross-validation process, we used grid search with some set parameter values that we chose, to see which combination gave us the best model results. When we trained the model with the best combination of parameters, the resulting MAE from the cross-validation was 6.1, which is a somewhat moderate to low value when compared to the range of the target value.

We must note that, unlike the previous section, the data standardisation did not improve the model's performance. The range of the target value is bigger than in the previous case as we didn't standardise, so it's normal for the MAE and RMSE results to be higher in this situation.

As for the performance on the unseen test data, the resulting MAE was 5.4, while the RMSE was 8.9. This means that the deviation of our predictions from the actual values relative to the range of the target variable is relatively small, which suggests that our predictions are fairly accurate.

We should note that even though a relationship between the categorical variables (authority and setting type) and the target variable might be present, the two categorical variables weren't included in the final model training. We did this because, when we added them to the input data, they didn't improve the model

performance and didn't give additional information beyond what the model captured from the total number of settings and the number of people in post.

Conclusion

This way of estimating will be the basis for future collections. This is the first time we've used machine learning to predict missing values, so our predictions will become more accurate as data quality improves.