CENSUS PROJECT REPORT

This report provides a comprehensive and detailed analysis of the population data of a moderately sized town, situated in between two major cities. The aim is to make insightful recommendations on what the local government should build on an unoccupied piece of land.

To ensure accurate recommendations, errors in the data such as missing (null) values, spurious entries and misleading information were treated or dropped where appropriate.

Furthermore, a well detailed visualization and analysis will be used to arrive at the recommendations in the latter part of the report.

DATA CLEANING PROCESS

To maximize accuracy, the dataset was cleaned, and a brief insight into the cleaning process performed has been summarized in subsequent paragraphs below.

The ages represented in floats were converted to integers and the empty string (‘ ’) was replaced with the mean age of all university students in the data.

For the Marital Status, 18 years has been set as the minimum legal age for a person to be married or enter civil partnership (Marriage and Civil Partnership (Minimum Age) Act 2022, c.28). All null values for ages less than 18 were changed to ‘N/A’.

Marital Status	count	mean	std	min	25%	50%	75%	max
Divorced	791.0	42.415929	17.389367	16.0	29.0	39.0	52.0	100.0
Married	2221.0	50.488969	16.751617	18.0	38.0	48.0	62.0	116.0
Single	2999.0	38.247749	14.176446	17.0	26.0	37.0	49.0	96.0
Widowed	387.0	67.041344	18.265127	18.0	65.5	71.0	78.0	105.0

As seen from the raw data above, some entries were under 18 years but stated they were divorced. This is highly misleading because people less than 18 years cannot be legally married, let alone, divorced (Marriage and Civil Partnership (Minimum Age) Act 2022, c.28). These entries were changed to ‘N/A’ and couples violating the government law were dropped in cases where their relationship could not be justified from the data information. Also, it is highly likely that people could be divorced, single, married or widowed at an old age.

Since the minimum marriageable age is 18 years, it is possible, although with a very low probability of occurrence, that an 18-year-old could get married and become widowed (Marriage and Civil Partnership (Minimum Age) Act 2022, c.28).

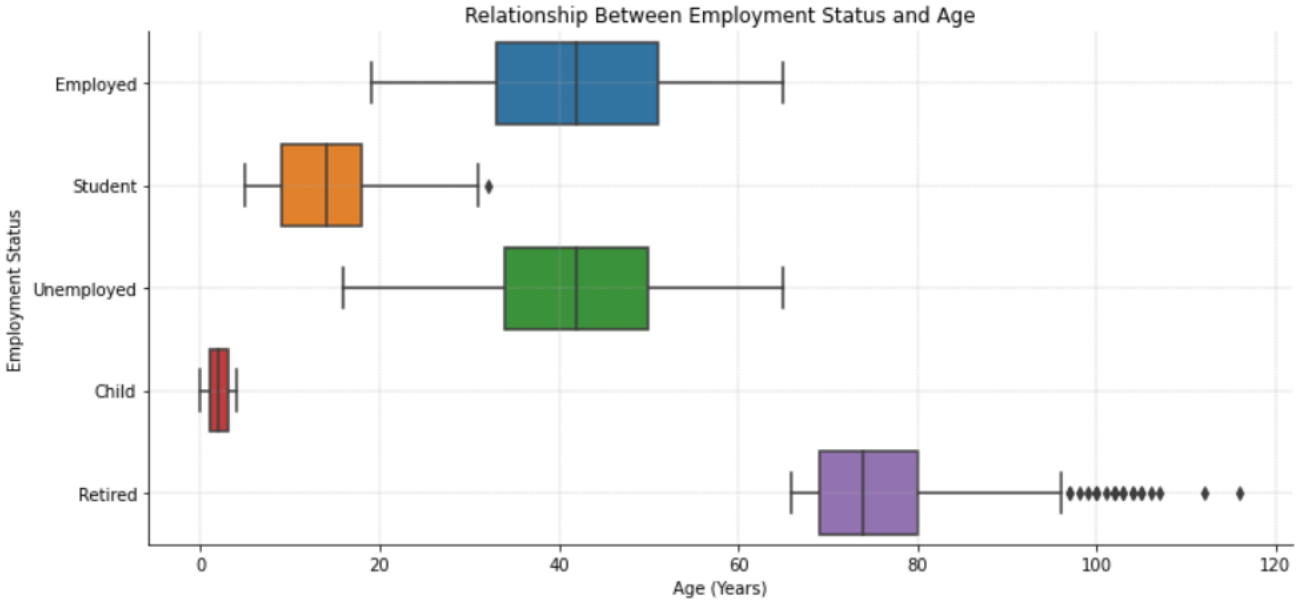
Marital Status	count	mean	std	min	25%	50%	75%	max
Divorced	790.0	42.449367	17.374919	18.0	29.0	39.0	52.0	100.0
Married	2221.0	50.488969	16.751617	18.0	38.0	48.0	62.0	116.0
N/A	1977.0	8.982296	5.072769	0.0	5.0	9.0	13.0	17.0
Single	2998.0	38.254837	14.173496	18.0	26.0	37.0	49.0	96.0
Widowed	387.0	67.041344	18.265127	18.0	65.5	71.0	78.0	105.0

The cleaned marital status feature now contains correctly assigned values as shown above.

Empty strings (values) in the data, for example Gender, Surname, Marital Status, etc, were treated by examining other entries in the data either living in the same household or bearing similar characteristics across other features. Some values, like first name, which require more data to determine, were filled with 'No Entry' or 'Undecided'.

Also, entries with wrong spelling ('Neice' in Relationship to Head of House) were corrected to the correct spelling.

The retirement age in the United Kingdom has been set to 66 years (Pensions Act 2014, s.26). Hence, all entries above 65 years, including those who entered a different occupation, were assigned 'Retired'. Also, all retired professions were group and assigned to 'Retired' as Occupation.



To aid analysis, 'Employment Status' feature was added to the data to classify the entries into employed, unemployed, students, child and retired as shown above.

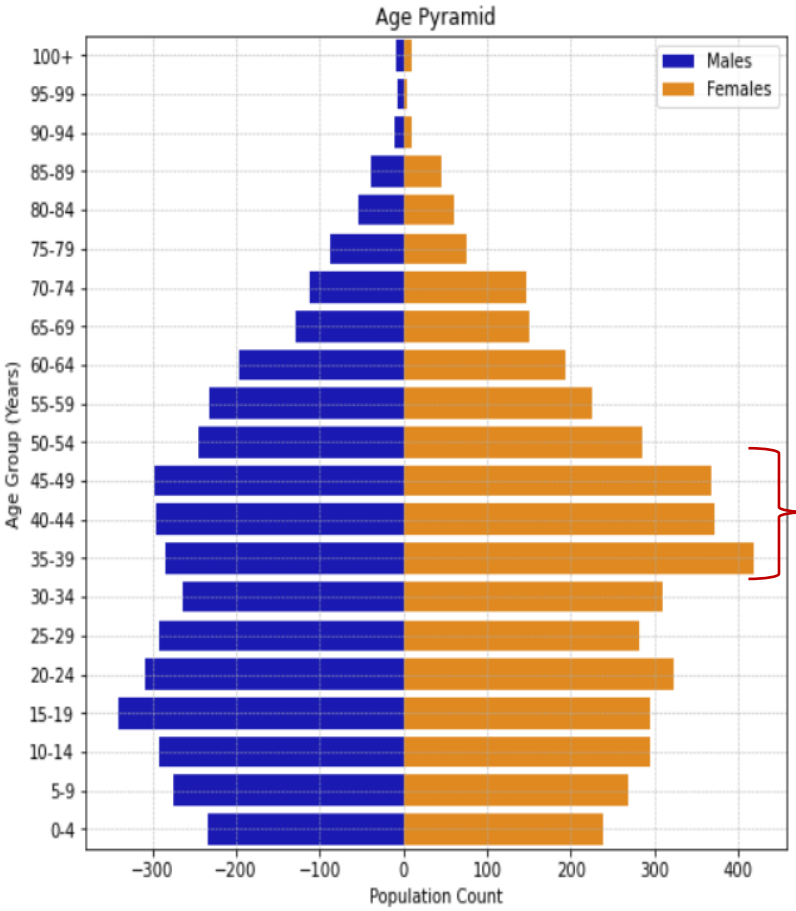
Undisclosed infirmity records (empty strings) were replaced with 'Undisclosed' since it couldn't be determined from the information contained in the data.

For religion, catholic and methodist were classified as Christian while private, orthodoxy, jedi and housekeeper were classified as 'Unidentified'. Jedi is regarded as misleading entry as it is not a recognized religion in the UK (Independent, 2016). Based on the assumption that ages under 18 are still minors, all missing values in religion for minors (less than 18 years) were replaced with 'Minor (Under 18)' while those above 18 years with missing religions were classified as 'Undecided'.

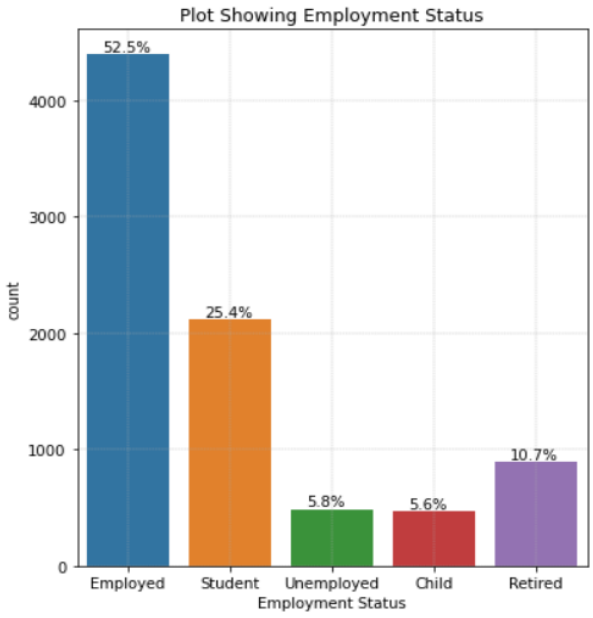
It is important to note that the outliers observed for the age feature e.g., above 100 years, were retained for the purpose of analysis. However, for future works such as model building, these outliers should be treated.

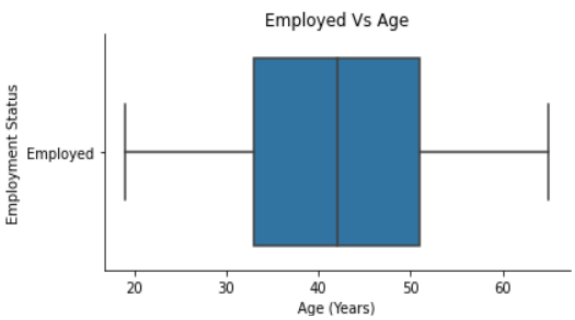
DETAILED ANALYSIS

Population (Age) Distribution



	Ages	Males	Females	Population Count
0	100+	-8	9	17
1	95-99	-6	4	10
2	90-94	-10	10	20
3	85-89	-38	45	83
4	80-84	-54	61	115
5	75-79	-87	75	162
6	70-74	-111	147	258
7	65-69	-129	151	280
8	60-64	-196	193	389
9	55-59	-232	225	457
10	50-54	-245	285	530
11	45-49	-298	368	666
12	40-44	-296	371	667
13	35-39	-285	419	704
14	30-34	-264	309	573
15	25-29	-291	281	572
16	20-24	-309	322	631
17	15-19	-341	295	636
18	10-14	-292	294	586
19	5-9	-275	269	544
20	0-4	-234	239	473



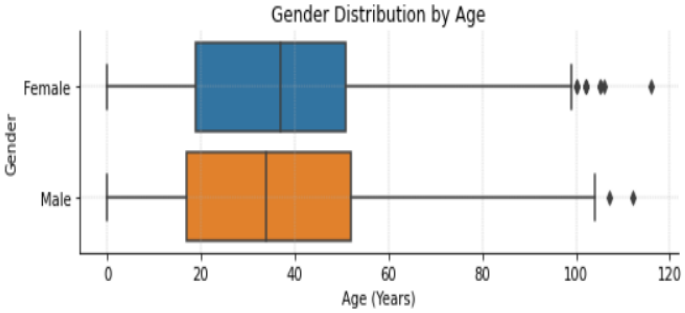


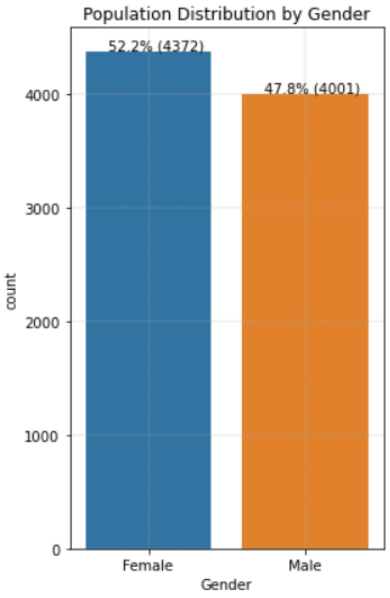
Although the population experienced a low birth rate (0-4), there is a steady growth into Middle Ages and then a decline in population from age range 45-49 as they grow into old age. Also, the population tend to live long into a good old age for both genders. This information will be crucial in calculating the birth and death rate for the population.

Furthermore, the bulge (growth) seen between age range 30-35 and 45-49, with population of 704, 667 and 666 respectively, could have been a result of people migrating to the town in search of employment – indicating job opportunities within the town. This is evidenced in the plot on the left with over 50% of population employed and between ages 32 and 51 years old.

Gender and Age

Gender	count	mean	std	min	25%	50%	75%	max
Female	4372.0	36.788884	21.539247	0.0	19.0	37.0	51.0	116.0
Male	4001.0	35.796301	22.036668	0.0	17.0	34.0	52.0	112.0

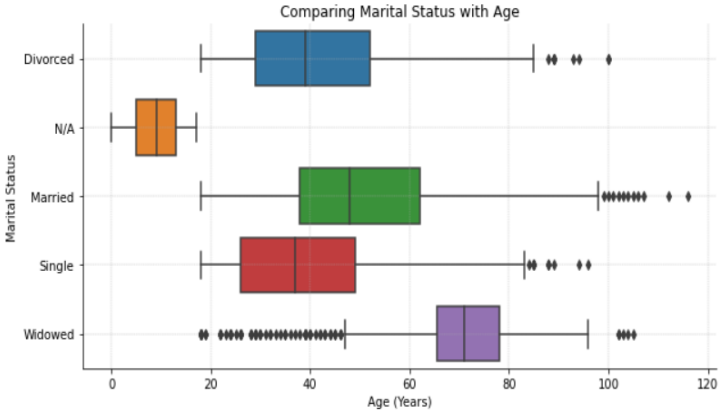


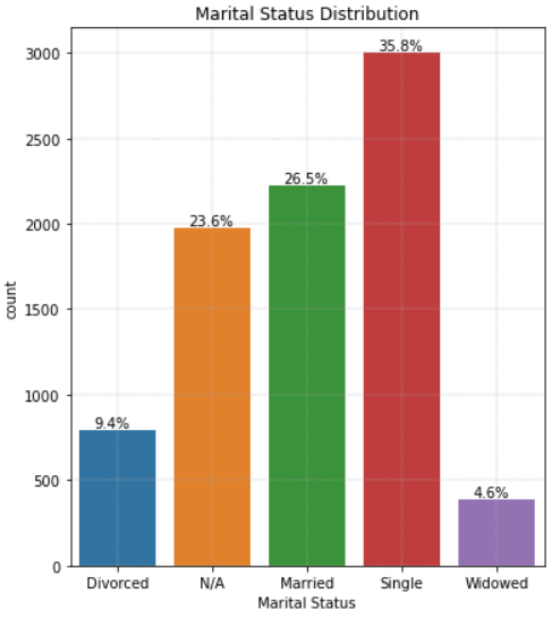


The number of females in the town is about 4.4% higher than that of males and a larger percent of both genders are concentrated between the ages of 17 to 52 and 19 to 51 years old for male and females respectively. Also, both genders live well into old age, as indicated by the outliers. This would come in handy when deciding the kind of care home to build, if need be.

Marital Status and Age

Marital Status	count	mean	std	min	25%	50%	75%	max
Divorced	790.0	42.449367	17.374919	18.0	29.0	39.0	52.0	100.0
Married	2221.0	50.488969	16.751617	18.0	38.0	48.0	62.0	116.0
N/A	1977.0	8.982296	5.072769	0.0	5.0	9.0	13.0	17.0
Single	2998.0	38.254837	14.173496	18.0	26.0	37.0	49.0	96.0
Widowed	387.0	67.041344	18.265127	18.0	65.5	71.0	78.0	105.0





To aid analysis, a new category 'N/A' was added to represent the marital status for minors while 'Single' represented those of legal marriageable age (above 17 years) who are yet to marry or enter a civil partnership (Marriage and Civil Partnership (Minimum Age) Act 2022, c.28).

About 4.6% of the town are widowed and the plot against age shows that younger couples tend to become widowed as much as the older ones – although, most of the widowed are between 65 and 78 years of age.

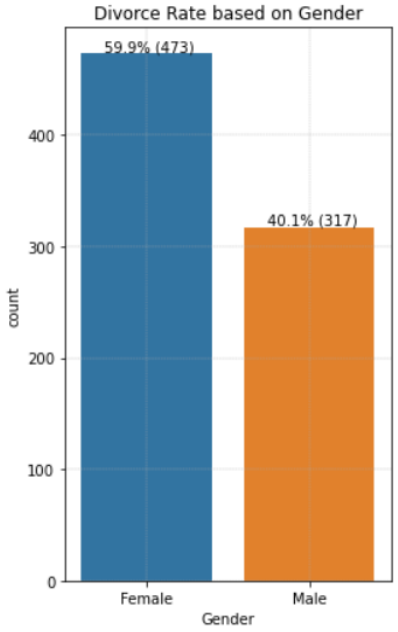
A larger portion of the population (35.8%) are Singles within the age of marriage (above 18 years) (Marriage and Civil Partnership (Minimum Age) Act 2022, c.28).

Marriage and Divorce Rate

$$Crude\ Marriage\ Rate = \frac{No.\ of\ Marriages}{Total\ Population}\ x\ 1000$$

$$Crude\ Divorce\ Rate = \frac{No.\ of\ Divorces}{Total\ Population}\ x\ 1000$$

Gender	Marital Status	count	mean	std	min	25%	50%	75%	max
Female	Divorced	473.0	43.640592	17.560483	18.0	31.0	40.0	53.0	100.0
	Married	1117.0	49.436885	16.701759	18.0	37.0	48.0	60.0	116.0
Male	Divorced	317.0	40.671924	16.966635	18.0	26.0	38.0	51.0	84.0
	Married	1103.0	51.577516	16.730736	18.0	39.0	50.0	63.0	112.0



In other to compute the marriage and divorce rates, the total number of marriages and divorces were set as the highest count between females and males. This was done due to the imbalance between the total number of married and divorced for both genders in the data. As a result, the total number of marriage and divorces was set to the total count for females (1117 and 473 respectively).

The crude marriage rate was calculated to be 133 marriages per thousand while the crude divorce rate was 56 divorces per thousand while the divorce to marriage rate was 42%. This indicates that the population tend to get married more than they get divorced.

Furthermore, from the figure on the left, divorced females (59.9%) are more than males (40.1%). A reason for this could be that some men left the town to settle somewhere else after getting divorced.

Birth and Death Rate

$$Crude\ Birth\ Rate = \frac{Number\ of\ Births}{Total\ Population}\ x\ 1000$$

$$Crude\ Death\ Rate = \frac{Number\ of\ Deaths}{Total\ Population}\ x\ 1000$$

To compute crude birth rate, it was assumed that entries with ages 0 years were new-born babies and total number across the data represents the number of births. Crude birth rate was found to be 11 births per thousand and the crude birth rate 5 years ago was also found to be 13 births per thousand (18% higher than now), indicating a low birth rate compared to 5 years ago.

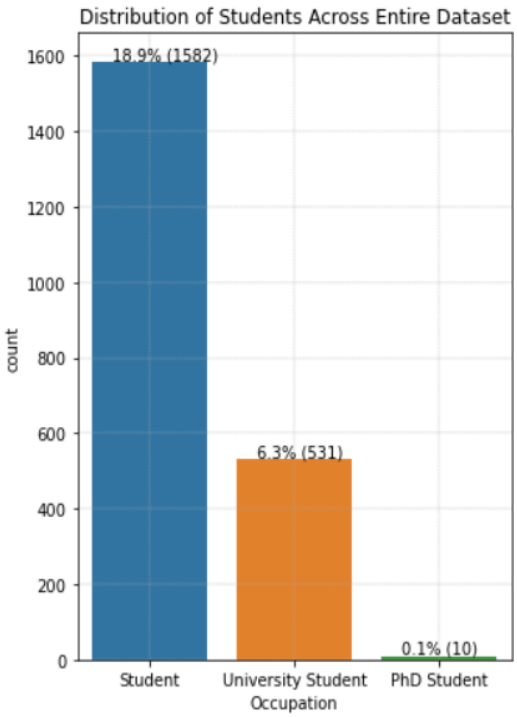
Similarly, to compute the death rate, it was assumed that only those beyond the retirement age remained in the town until they died while others still within employment age could have left the town seeking opportunities elsewhere, hence the reduction in population from 65 years up, shown by the minus (-) in cumulative count. The death rate was then calculated to be 6.5 deaths per thousand by obtaining the average cumulative count for ages above 65 years.

	Ages	Population Count	Cummulative Count
0	100+	17	7
1	95-99	10	-10
2	90-94	20	-63
3	85-89	83	-32
4	80-84	115	-47
5	75-79	162	-96
6	70-74	258	-22
7	65-69	280	-109
8	60-64	389	-68
9	55-59	457	-73
10	50-54	530	-136
11	45-49	666	-1
12	40-44	667	-37
13	35-39	704	131
14	30-34	573	1
15	25-29	572	-59
16	20-24	631	-5
17	15-19	636	50
18	10-14	586	42
19	5-9	544	71
20	0-4	473	0

Commuters

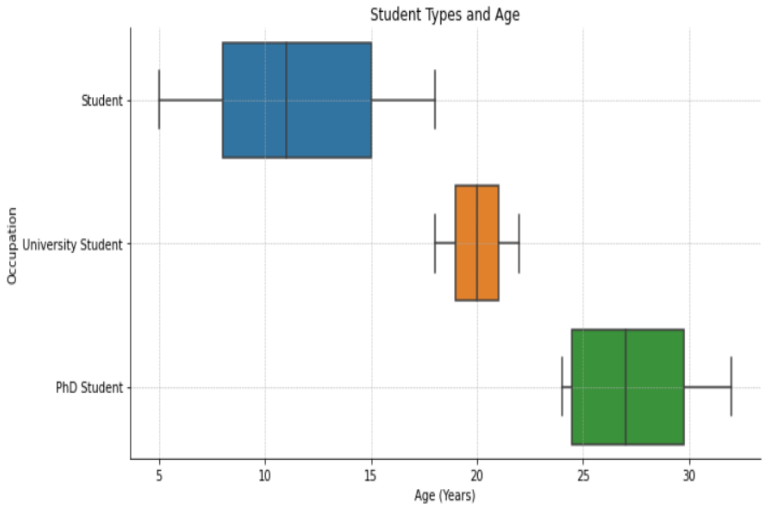
For ease of analysis, three major points were used to estimate number of commuters in the town:

- The total number of students attending universities (University and PhD Students).
- The number of university staffs (all entries with lecturer or higher education included in their occupation title) – higher education refers to tertiary institutions (National Career Services, no date).
- Non-tertiary students (above 8 years) who would be of university age in 10 years' time.



Occupation	count	mean	std	min	25%	50%	75%	max
PhD Student	10.0	27.400000	2.988868	24.0	24.5	27.0	29.75	32.0
Student	1582.0	11.481669	3.900321	5.0	8.0	11.0	15.00	18.0
University Student	531.0	20.327684	1.221647	18.0	19.0	20.0	21.00	22.0

75% (Above 8 years)



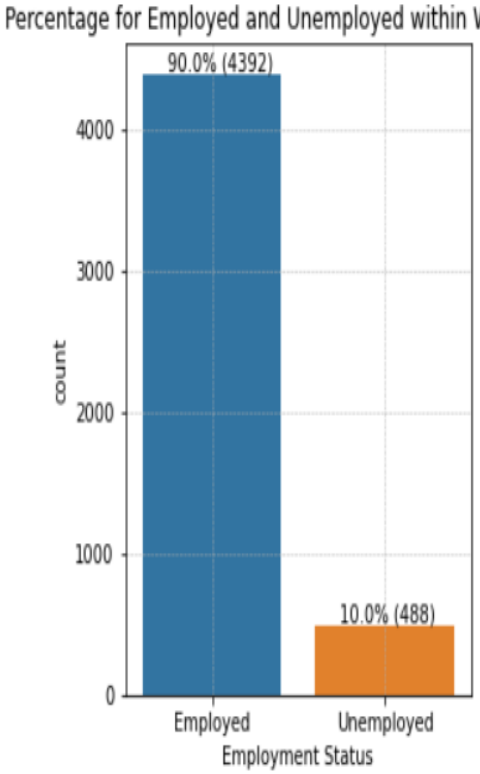
The students currently attending universities (both university and PhD students) outside the town accounts for a total of 6.4% (541) of the total population (8373).

According to the data information, there were no universities in the town, hence all students in higher education commutes. So, it was assumed that all higher education workers (lecturers, career advisers, etc) were the major commuters outside the town. Total number of employed commuters was obtained by summing the number of higher education workers, which gave 38 commuters.

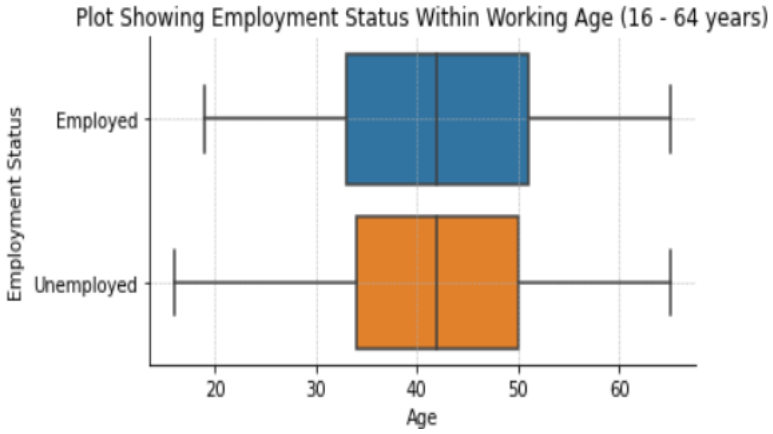
Overall, there are a total of 579 commuters in the town, with higher education students (University and PhD students) as the majority.

Furthermore, analysing the non-university students currently under 18 years, 75% of these students are between 8 and 18 years old. This means that in at least 10 years from now, these students would have attained university age range as specified by the minimum age of university student (18 years – from the minimum age of university students in the town) and will join the commuters for studies at universities. Hence, a pointer to the fact that train stations should be built to decongest the road networks both for current and futuristic benefits.

Employment and Unemployment Trend



Employment Status	count	mean	std	min	25%	50%	75%	max
Employed	4392.0	42.168488	11.692672	19.0	33.0	42.0	51.0	65.0
Unemployed	488.0	42.487705	11.260763	16.0	34.0	42.0	50.0	65.0

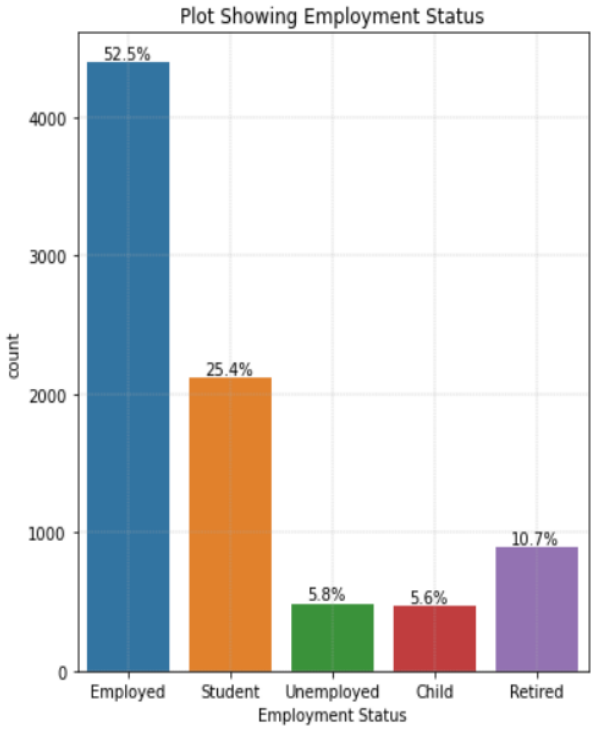


From the plot above, 4392 (90%) of the population within working age (16 to 65 years) are employed while 488 (10%) are still unemployed – this is excluding students in higher education who are of working age (Legal Solicitors, no date). Although the number of those employed supersedes the unemployed, it is worthwhile to pay attention to the unemployment rate because previous analysis of the age pyramid indicated an influx of people between age range 35-39 and 45-49 and from the illustrations above, about 50% of those unemployed within this range. This could mean a fraction of people migrating to the town are unemployed.

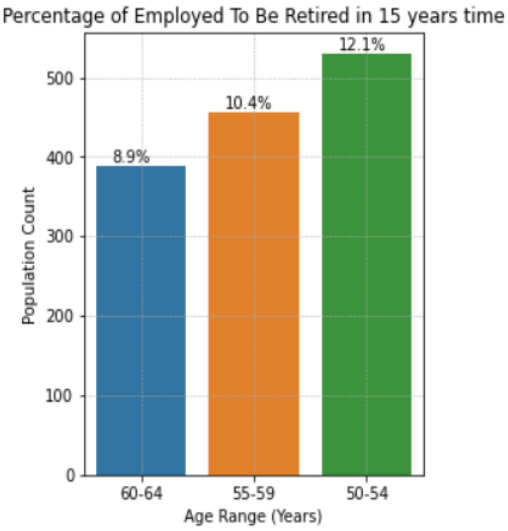
Occupation	count	mean	std	min	25%	50%	75%	max
PhD Student	10.0	27.400000	2.988868	24.0	24.5	27.0	29.75	32.0
Student	1582.0	11.481669	3.900321	5.0	8.0	11.0	15.00	18.0
University Student	531.0	20.327684	1.221647	18.0	19.0	20.0	21.00	22.0

In addition, the total number of students currently enrolled in higher education (university and PhD. Students) is 541 (about 6.4% of the entire population). This means that, in few years to come, these students currently studying at the university would be ready for full employment and this could have an impact on unemployment rate as there is also an influx of people into the town looking for employment.

Future Retirement

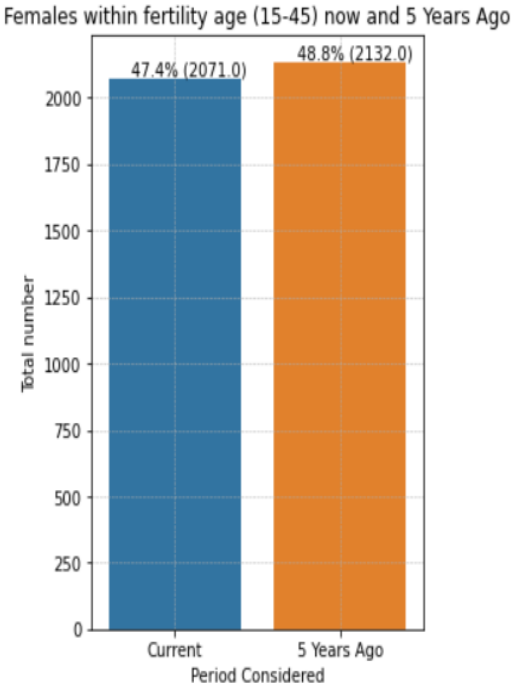


	Ages	Males	Females	Population Count
8	60-64	-196	193	389
9	55-59	-232	225	457
10	50-54	-245	285	530



Slightly over 10% of the entire population are retired and based on an assumption that those working stay and retire in the town, approximately 31.4% of those employed will be in retirement 15 years from now, thereby, adding to the number of those retired. The 31.4% is derived by summing the percentage of the town currently that are working and are, at most, 15 years from retirement (between 50 and 64 years) as shown in the plot on the right above. This will be key in deciding whether the authorities should prepare for this cater for them or not.

Fertility Rate



The overall childbearing age range for women has been assumed to be between 15 and 45, although there are instances (but very few) where a woman could give birth beyond this age (Office for National Statistics, 2022).

By setting the fertility age at range of 15 to 45, the number of females within this age bracket was 2071 (47.4% of the entire female population) and 2132 (48.8% of female population) 5 years ago. This means there was a 2.9% drop in number of females within childbearing age which further explains the low birth rate observed from the base of the age pyramid.

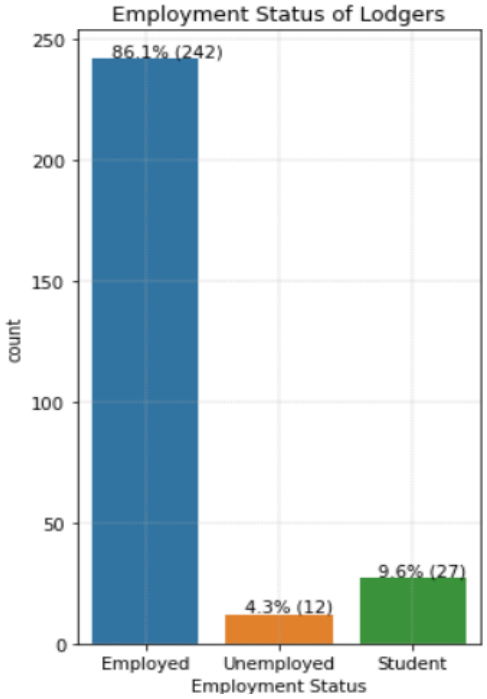
IMMIGRATION, EMIGRATION AND NET MIGRATION RATES

$$Immigration\ rate(IR) = \frac{number\ of\ immigrants}{Total\ Population}\ x\ 1000$$

$$Emigration\ rate(ER) = \frac{number\ of\ emigrants}{Total\ Population}\ x\ 1000$$

Employment Status	count	mean	std	min	25%	50%	75%	max
Employed	242.0	41.789256	12.504979	19.0	30.0	41.0	52.00	64.0
Student	27.0	20.666667	1.270978	18.0	20.0	21.0	22.00	22.0
Unemployed	12.0	39.166667	12.597499	24.0	29.5	37.0	46.25	63.0

$$Net\ migration\ rate = IR - ER$$

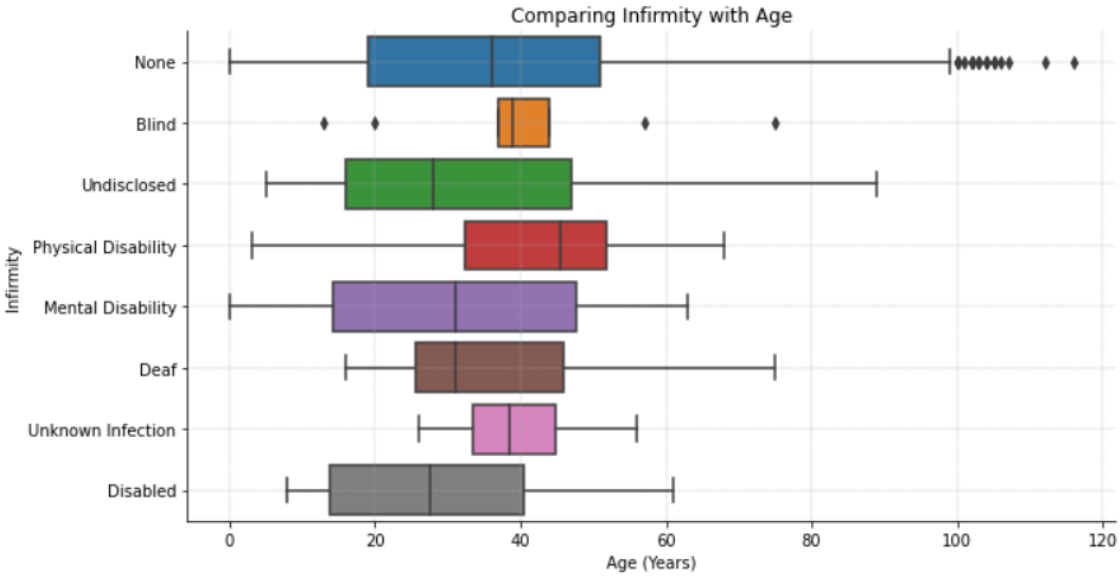


To compute the immigration and emigration rate, the number of immigrants has been assumed to be the number of lodgers in the town while the number of emigrants assumed to be the difference between the female and male divorcees – since the former is greater than the latter (Eurostat, 2018).

Immigration, emigration and net migration rate was found to be 34, 19 and 15 migrants per thousand respectively.

This means that more people enter more than they leave the town and from the plot, over 85% of these immigrants (lodgers) are employed with more of them between 30 and 52 years. This suggests that most people migrate into the town for employment.

Infirmity Status



Although there are cases of infirmity across all ages, people above 85 years suffered no infirmity which could mean that they are well taken care of at old age.

RECOMMENDATIONS

Coupled with the estimated number of commuters currently living in the town, there seem to be a good number of students (non-tertiary) who would have attained the minimum age required by law to enter the university in approximately ten years' time. Hence, it is imperative for the local authorities to build train and subway stations to prepare for this increased number of commuters in the future. The analysis also suggests an influx of people within the working age into the town and this may increase as the town further develops. So having a train station to cater for the influx of people into the town, as well as commuters leaving the town, would efficiently decongest the road network.

Although the population within the employment age range had a good percentage employed, there were still a few who were unemployed (ages predominantly between 34 and 50 years). Also, it is worthwhile to note that students in higher institutions would soon enter the job market and would have to compete with those currently unemployed as well as others from outside the town seeking jobs within the town. Hence, an investment in employment and skill building would highly impact positively in further reducing the unemployment rate and ensure most university graduates are employable.

Furthermore, in addition to investment in infrastructure due to influx of people into the town in search of job or settlement, there would also be a need to invest in care as seen from the future retirement analysis because there is likely going to be over 30% increase in the number of the population above retirement age and may be needing care in the future, even though the population appear healthy at old age.

REFERENCES

Pension Act (2014) Section 26
Available online: https://www.legislation.gov.uk/ukpga/2014/19/section/26 [Accessed 15/04/2023]

Marriage and Civil Partnership (Minimum Age) Act (2022) chapter 28
Available online: https://www.legislation.gov.uk/ukpga/2022/28/enacted [Accessed 15/04/2023]

Independent (2016) *Jediism is not a religion, Charity Commission says*
Available online: https://www.independent.co.uk/news/uk/home-news/jediism-star-wars-not-religion-charity-commission-a7484856.html [Accessed 10/04/2024]

World Health Organization (No Date) *Crude birth rate*
Accessed online: https://www.who.int/data/gho/indicator-metadata-registry/imr-details/1139 [Accessed 15/04/2023]

National Careers Service (No Date) *Higher Education*
Accessed online: https://nationalcareers.service.gov.uk/explore-your-education-and-training-choices/higher-education [Accessed 18/04/2023]

Legal Solicitors (No Date) *What is the Minimum Legal Age of Employment in the UK?*
Available online: https://www.localsolicitors.com/employment-guides/what-is-the-minimum-legal-age-of-employment-in-the-uk [Accessed 18/04/2023]

Office for National Statistics (2022) *Childbearing for women born in different years, England and Wales: 2020*
Available online:
https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/bulletins/childbearingforwomenbornindifferentyearsenglandandwales/2020#:~:text=A%20woman's%20childbearing%20is%20assumed,not%20affect%20the%20overall%20patterns [Accessed 20/04/2023]

Eurostat (2018) *Glossary:Migration*
Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Migration
[Accessed 22/04/2023]