Census Report

This study gives a thorough overview of a town with 10173 residents. It describes the town's living circumstances in some depth and goes into detail about the residents' experiences considering the census data. The recommendations made using the data provided by the individuals will assist to improve the quality of life in the community, by developing unused land in the municipality. The report is divided into the following parts: Cleaning the data to get rid of errors that were either intentionally or unintentionally captured. The report will then emphasise a few crucial conclusions drawn from the data before providing recommendations.

Data Cleaning.

The Jupyter notebook that is linked to this report details the process that was used to carefully inspect the census data for inaccuracies. Most of the missing entries were filled in by determining a person's home, finding their entry number, and updating data (in the case of the missing street and household number). A family without a surname and an entry without a first name were both substituted with the highest frequent surname found in the data. In the case of Religion, when the entry was discovered to be over 18, but the entries were "undecided," "na," or "Sith," insignificant logs that could not be inferred from the household were substituted with "None" (BBC, 2016) Most entries for people under the age of 18 seemed to have "Nan" entries, which were later changed to "Underaged" pending the time when they could make fully informed decisions (18), however a handful already identified as Christians. Even though some of them may have inherited their parents' faith, this does not mean that their beliefs are infallible. Similar to marital status, entries below the age of 18 are considered to be minors while those above the age of 18 are considered to be single. There was a record that was thought to be inaccurate because the person listed lived alone, was not connected to a family, was a minor, but after inferring from the data he was attached to a family based on his location.

Population Demographics

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 10174 entries, 0 to 10173
Data columns (total 14 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   House Number                    10174 non-null  int32
 1   Street                          10174 non-null  object
 2   First Name                      10174 non-null  object
 3   Surname                         10174 non-null  object
 4   Age                             10174 non-null  int32
 5   Relationship to Head of House   10174 non-null  object
 6   Marital Status                  10174 non-null  object
 7   Gender                          10174 non-null  object
 8   Occupation                      10174 non-null  object
 9   Infirmity                       10174 non-null  object
 10  Religion                        10174 non-null  object
 11  Occupation_Bar                  10174 non-null  object
 12  Age_column                      10174 non-null  object
 13  Occupancy Count                 10174 non-null  int64
dtypes: int32(2), int64(1), object(11)
memory usage: 1.1+ MB
```

Table 1 Data columns and attributes

The census data has the following characteristics after being cleansed.

The data were further divided into the following categories to facilitate in-depth analysis: • Age column: Ages classified into 5-year age groups for the population's age pyramid. • Occupation Bar: Breaks down the population of the town into groups such as children, students, Ph.D. students, university students, and employed people.

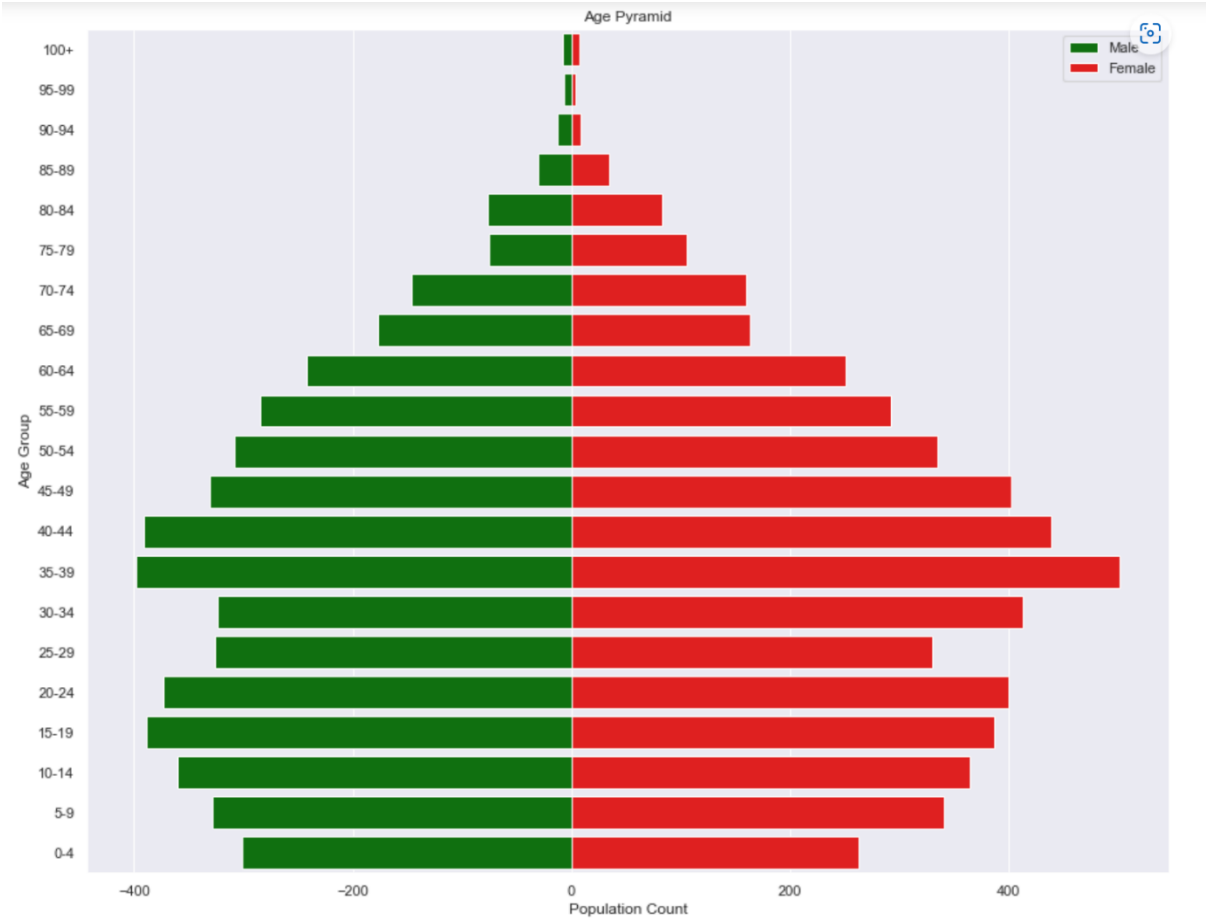

Figure 1 Age Pyramid of the Population

Marital Status

Value	Count	Frequency (%)
Single	3554	34.9%
Married	2893	28.4%
Minors	2416	23.7%
Divorced	910	8.9%
Widowed	400	3.9%

Table 2 Marital Status

Occupation

Value	Count	Frequency (%)
Employed	5431	53.4%
Student	1982	19.5%
Retired	856	8.4%
Unemployed	706	6.9%
University Student	618	6.1%
Child	565	5.6%
Phd Student	15	0.1%

Table 3 Occupation statistics

Religion

Value	Count	Frequency (%)
None	3426	33.7%
Underaged	2466	24.2%
Christian	2311	22.7%
Catholic	1061	10.4%
Methodist	666	6.5%
Muslim	146	1.4%
Sikh	54	0.5%
Jewish	35	0.3%
Orthodoxy	3	< 0.1%
Bahai	2	< 0.1%
Other values (3)	3	< 0.1%

Table 4 Religion

Infirmity

Value	Count	Frequency (%)
None	10100	99.3%
Physical Disability	18	0.2%
Disabled	15	0.1%
Blind	12	0.1%
Unknown Infection	11	0.1%
Deaf	9	0.1%
Mental Disability	8	0.1%

Table 5 Infirmity Count

According to additional descriptive research, the town's employment rate is over 50% as shown *table 3*, with a large proportion of middle-aged individuals and 7% of the population being unemployed. The majority of people in the community are either married or single as seen in *table 2*, and infirmity is minimal when compared to other factors as seen in *table 5*. The population appeared to be largely healthy, and while other records indicated the presence of an unidentified infection, which might indicate the beginning of a serious illness in the town, we can't make any significant conclusions from it because we don't have enough information. There will therefore be no recommendations on the effects of illness for this town because no more information has been given.

Outliers and Deviations.

	Marital Status	count	mean	std	min	25%	50%	75%	max
0	Single	3554	38.15	18.21	18	26	37	49	88
1	Married	2893	50.36	16.77	18	38	49	62	109
2	Divorced	910	41.99	17.05	18	28	39	53	94
3	Widowed	400	64.87	19.06	18	65	70	77	103

Table 6 Marital status metrics

A significant portion of young people are widowed, according to the data from table 6. Further research reveals that most of these entries are university students, the majority of whom can be presumed to be immigrants, and that this appears very improbable. As a result, we will categorise this information as being false or being too unclear to be useful for our study. Single, Married, and Divorced people above the age of 80 are not unusual.

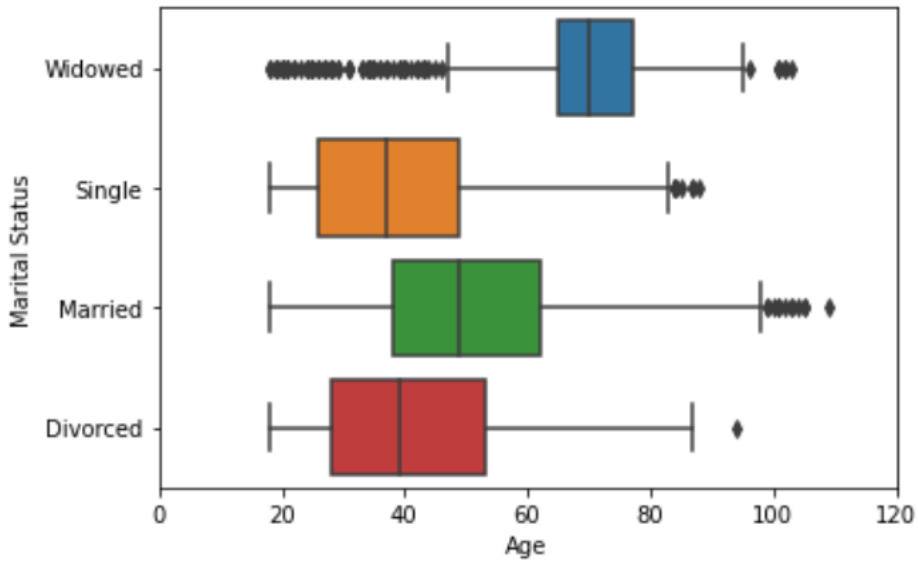

Figure 2 Marital Status based on Age of population

This whisker plot illustrates the employment level in the data, which were grouped based on the census.

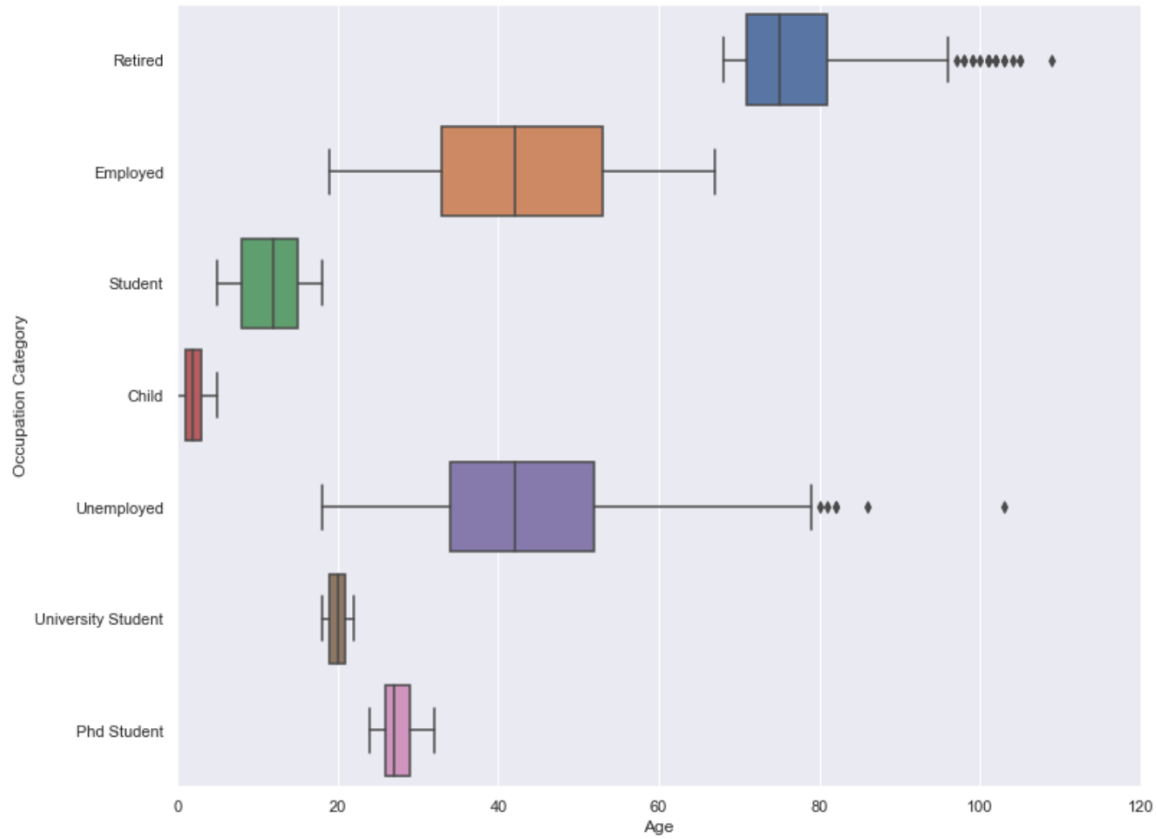

Figure 3 Employment Ratio in the Population

A few individuals over the age of 79 who were listed as unemployed are not unusual given that the UK has endorsed the right of everyone to work until they choose to retire (Age UK, 2021), over-80 retirees are also not a cause of concern. Figure 4 gives a deeper look at the gender disparity in unemployment ratio; the town has more unemployed women than men. Within the age 30 to 50, unemployment is at its peak.

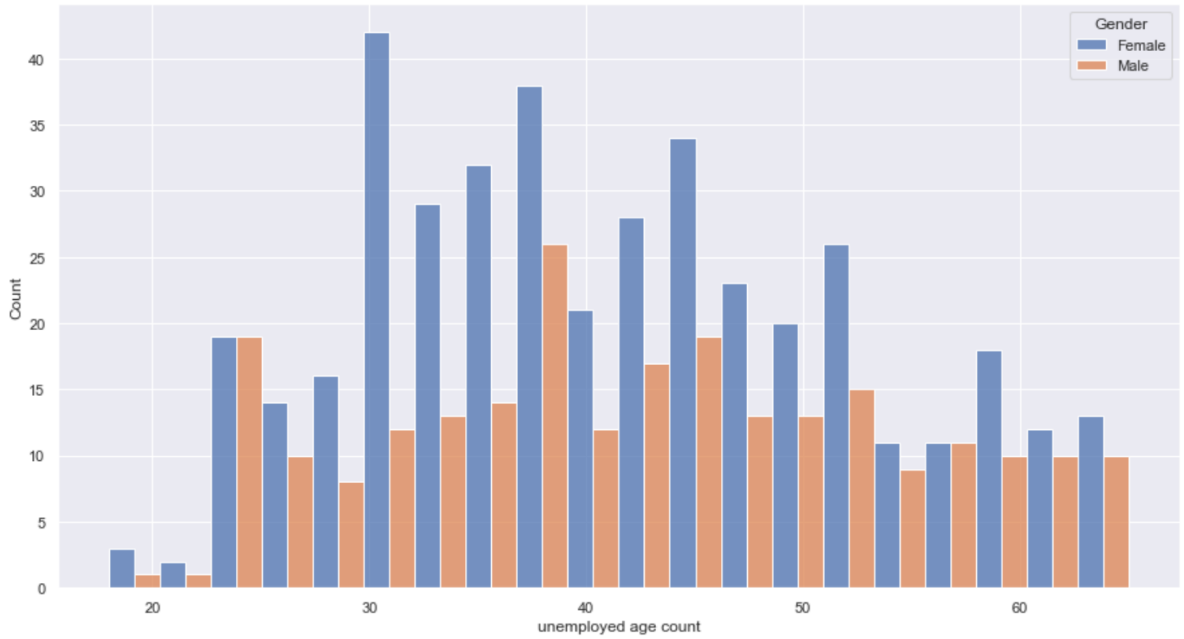

Figure 4 Unemployment ratio based on Gender

Comparing the number of economically active residents in the municipality to the number of unemployed persons yielded a crude unemployment rate of just 9%. This might be regarded as being low, yet (D. Clark, 2021) indicated in his report that, as of September 2021, the UK had a really high unemployment rate of 4.3%. This significantly deviates from our 9% number, leading us to believe that unemployment may be a problem in this community.

Religion

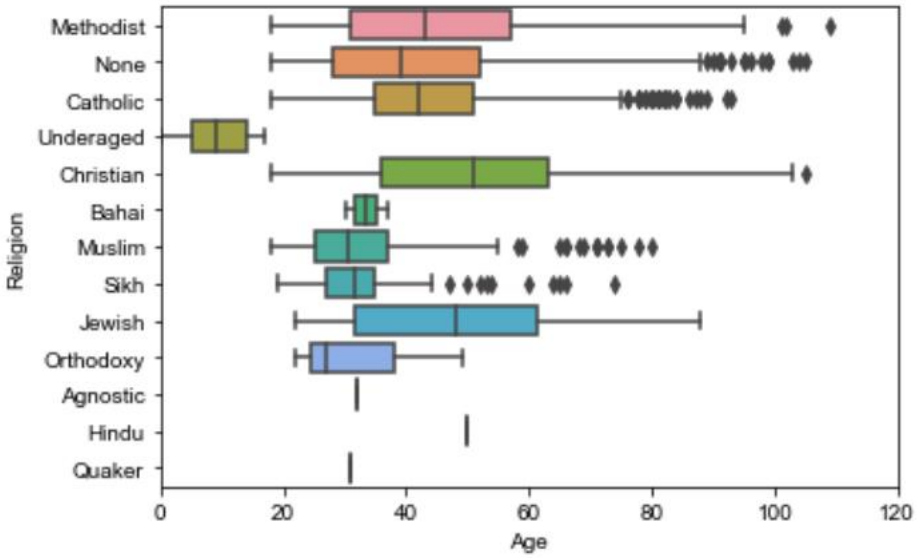

Figure 5 Religion Ratio of the population

Table 4 shows that, in contrast to the many other religions listed, 33.7% of the town is irreligious. Most religions are expanding although they only make up a little portion of the population.

Religion	mean	median
Agnostic	32.000000	32.0
Bahai	33.500000	33.5
Catholic	43.734213	42.0
Christian	49.940286	51.0
Hindu	50.000000	50.0
Jewish	47.228571	48.0
Methodist	44.923423	43.0
Muslim	34.520548	30.5
None	42.031524	39.0
Orthodoxy	32.666667	27.0
Quaker	31.000000	31.0
Sikh	34.648148	31.5
Underaged	9.248175	10.0

Table 7 Religion Metric

Additionally, Christians made up 22.7% of the population, which is much more than any other religion represented in the municipality, with a mean age of 51. The fact that there is a Catholic church in the town does not serve as a place of worship for all Christians, but it also does not prove that a church is absolutely necessary for the community.

Divorce and Marriage

We can agree that not all marriages last forever. Table 6 shows that the minimum age for divorce was 18, which is rather early and can be due to a lot of complications such as early marriages and distance. There are several ways to calculate divorce rates, but we will be considering the crude divorce rate.

$$Crude\ Divorce\ Rate\ = \frac{total\ number\ of\ divorced\ couples}{Total\ Population} * 1000$$

It is inferred from the calculation above that there are 89 divorces per 1000 residents according to the census statistics. There are more female divorcees than male divorcees, which may indicate that more male divorcees depart the community.

Marital Status	Gender	
Divorced	Female	42.366920
	Male	41.489583
Married	Female	49.404255
	Male	51.362561
Minors	Female	9.192982
	Male	8.945082
Single	Female	38.611853
	Male	37.613095
Widowed	Female	64.834783
	Male	64.923529

Name: Age, dtype: float64

Table 8 Marital Status based on Gender

Population Growth Rate.

The crude birth rate (CBR) and crude death rate (CDR) are statistical measures to estimate the birth rate and death rates in a population calculated

thus,

$$CBR\ = \frac{Number\ of\ births\ * 1000}{Estimated\ Population\ at\ midyear}$$

$$CDR\ = \frac{Number\ of\ deaths\ * 1000}{Estimated\ Population\ at\ midyear}$$

After calculating the birth rate, this town's current crude birth rate is 10 births per thousand. This will be termed as low, based on the Birth Rate by Country (2021) it was postulated that the average global birth rate is 18.5 births per 1,000 total populations setting our population to be way below average. Five years prior, the CBR was estimated at 13 births per thousand. This shows that it has receded within the 5-year span. The Crude Death Rate can be defined to be the number of deaths in an area, the death rate for this town was calculated by estimating death based on the difference in the number of people within each age column for people over 65. This is because natural deaths are more inherent within this age group. The decline within the other age columns doesn't explicitly explain death as there's a huge possibility of migration being the leading cause in the lower age group. Thus, by summing the absolute value of the difference between the estimated amount of people in the 65-age group in the table and dividing by 5 (this accounts for the age grouping). We infer that the crude death rate is 13 deaths in a thousand residents.

Age_column	Count
95-99	-11.0
90-94	-44.0
85-89	-94.0
80-84	-21.0
75-79	-125.0
70-74	-34.0
65-69	-153.0
60-64	-83.0
55-59	-67.0
50-54	-25.0
5-9	-65.0
45-49	-97.0
40-44	-69.0
35-39	162.0
30-34	81.0
25-29	-117.0
20-24	-3.0
15-19	761.0
100 and Above	-709.0
10-14	160.0
0-4	NaN

Table 9 Death Rate over the Age groups

Migration

Migration is a constant throughout the world; it is defined as people migrating into and out of a place in search of a better life or in an effort to save money. In Figure 1, there appears to be a significant difference between the age groups of 20 to 24 and 25 to 29. The age range depicted in figure 3 primarily represents university and doctoral students who moved to the area. These individuals cannot be regarded as a permanent component of the population since we cannot predict where they will choose to live once their programmes are complete. In addition, new students will arrive in succeeding years to fill the vacant apartments. Hence, lodgers and single tourists were considered as immigrants. Many divorced lodgers resided in the town with their husbands before the split, so they were also not taken into account. The inference may be incorrect, but a rough estimate of the entire population is 26 immigrants per thousand headcounts.

$$Immigrants = \frac{Single\ lodgers}{Total\ Population} * 1000$$

We can also see that there are more female divorcees living in the town than male divorcees when emigrants are estimated by the difference between male and female divorces as shown in table 8.

Calculating the number of emigrants in the data based on the difference between them, gives an approximation of 13 emigrants per 1,000 headcounts. This was calculated by dividing the estimated number of emigrants by the total population in thousands. The immigrants make a double portion of the emigrants.

Commuters.

The census data's Occupation column was iterated through in order to find commuters. University students, Ph.D. candidates, and professors in higher education were all considered commuters. Teachers (except for those in higher education) and community workers were referred as non-commuters. To compare the commuters per thousand to the non-commuters, a rough estimate was made. According to this calculation, there were 129 commuters per 1,000 people, or 13% of the entire population.

Household Occupancy

This was carried out to estimate the number of occupants per household. The median is 2, and the mean is roughly above 2.

```
mean      2.791712
median    2.000000
var       3.107030
std       1.762677
Name: Occupancy_Bar, dtype: float64
```

Table 10 Household occupancy Metric

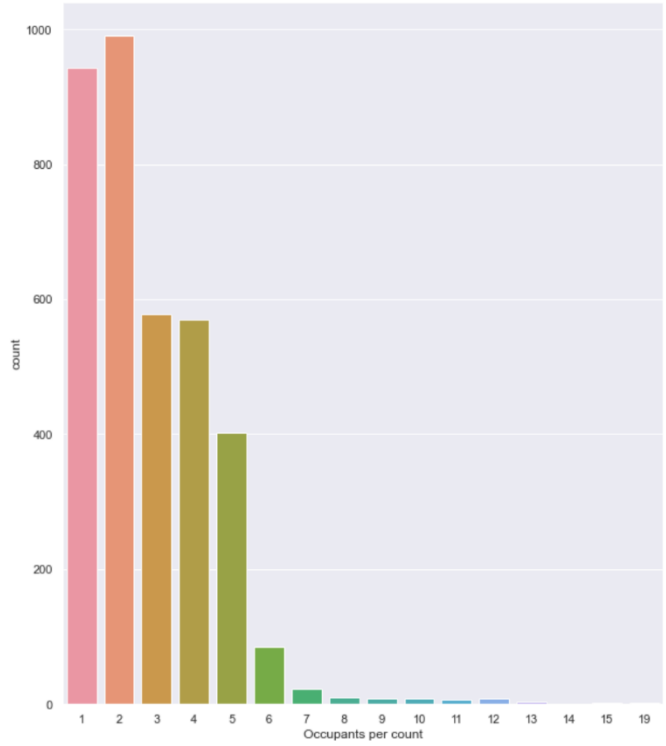

Figure 6 Household Occupants

The variance shows that the occupancy count is equally spread out in the town, meaning there is a healthy spread of high-density houses and low-density houses in the town. Identifying our median as 2, there are 1710 houses currently over-occupied.

	Occupancy Count	Age
0	1	72
1	2	31
2	3	37
3	4	39
4	5	8
5	6	20
6	7	7
7	8	13
8	9	20
9	10	5
10	11	21
11	12	14
12	13	12
13	14	1
14	15	65
15	19	25

Table 11 Modal Age based on occupancy

We can see the frequency of the occupant count across the modal age from the table 11. Due to the tendency of both young people and the elderly to congregate in one flat, this table demonstrates that occupancy is not exclusive to age. This can be because there are care facilities and orphanages nearby. Figure 6 shows high occupancy rates of 3, 4, and 5 individuals per apartment, with a modal age of 37, 39, and 8 for each bar, respectively. This might be explained by the town's apartment supply. This might be explained by the town's apartment supply. The average age of singles is 38.17 (table 6), which means that the majority of persons in that age group are single and share an apartment to cut down on mortgage expenses or because there aren't many apartments in the area. The 8-year-old children may be a result of the town's foster care facilities. Given that most males leave the town, and most women are unemployed, the high divorce rate in the population may be to blame. As a result, the majority of children are uprooted and placed in foster care, as seen by the occupancy rates of 14, 10, and 7. This doesn’t explain that the town is affluent because obviously, we don’t see their income, but it tells us that overcrowding seems to cut across all ages within the population.

RECOMMENDATIONS

Considering that there are more Christians than people of other religions in the town, it could seem necessary for them to have a place of worship, but not everyone in the community would benefit from this. As more people are expected to be displaced or forced to search outside of the community for skilled work in the future, unemployment is another pressing issue that should be given a thought to. As more males leave the town following divorce—which may also be due to divorce or a lack of employment—the population is already declining, hence many children are subsequently placed in foster care. The town's infirmity count was extremely low, but there seems to be an unidentified infection that could potentially be a growing pandemic because it hasn't been thoroughly investigated, a hospital may seem like a wise investment, and this can also aid the town's low birth rate problem. As most of these lodgers are unmarried and want to save money, overcrowding in flats may result from population expansion. Additionally, it looks that there is already a care home and a foster care home inside our population. The community will gain from high-density housing, as will some divorced families with small children and new lodgers. Although building a train station would seem like a risky investment, as opposed high-density housing that serves a larger portion of the population. Given the large number of commuters coming into and leaving the town frequently, it is very likely that a train station is advantageous to the community. More individuals will be willing to migrate into and out of the town due to the ease of access, which can encourage population expansion. This may also result in employment prospects as more people will be eager to relocate in search of employment. Additionally, investors may wish to make investments in this community, which will result in fewer family relocations.

Bibliography

Age UK (April 2021) Retirement Age. Available Online: Can I be forced to retire? | Work & Learning | Age UK [Accessed 30/11/2021]

BBC (2016) Jedi is not a religion, Charity Commission rules. Available online: Jedi is not a religion, Charity Commission rules - BBC News [Accessed 05/12/2021]

Birth Rate by Country 2021 Available Online: Birth Rate by Country 2021 (worldpopulationreview.com) [Accessed 08/12/2021]

Office for National Statistics (2020) Labour market in the regions of the UK: July 2020 Available online: A guide to labour market statistics - Office for National Statistics (ons.gov.uk) [Accessed 07/12/2021]

Statista (2021) Unemployment rate of the UK 1971-2021 Available online: UK: unemployment rate 2000-2018 | Statista [Accessed 05/12/2021]

UNCRC Article 14 Available Online: UNCRC Article 14 - I have the right to have my own thoughts and beliefs and to choose my religion with my parents' guidance - The Children and Young People's Commissioner Scotland (cypcs.org.uk) [Accessed 05/12/2021]