CENSUS REPORT

This report presents the findings of a census conducted in a moderately -sized town to provide recommendations for the development of an unoccupied plot of land, as well as to determine which services should be prioritized for investment by the local government.

The first part of this project has been provided where a thorough review of the census data cleaning has been undertaken. The analysis focuses on a range of demographic factors, including age pyramid, unemployment trends, religious affiliations, divorce and marriage rates, occupancy level, commuters, birthrate and death rate, population change for the town. These analyses have been designed specifically to support the recommendations provided in this report.

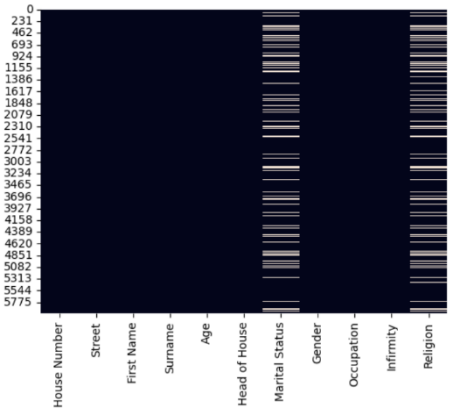
Summary of Numeric Data (Age)

The town's population is 6,000 individuals with a mean age of 37 and a standard deviation of 21.72, showing significant age variability. Age range spans from 0 to 105 years old. 25% of the population is 19 years or younger, 50% is 35 years or younger, and 75% is 52 years or younger.

	Age
count	6000.000000
mean	36.802667
std	21.722002
min	0.000000
25%	19.000000
50%	35.000000
75%	52.000000
max	105.000000

Data Cleaning and Imputation

The dataset has 6000 entries and 11 columns, with missing values in the marital status and religion columns. Various methods will be used to handle these missing values, with details available in the attached jupyter notebook.




```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6000 entries, 0 to 5999
Data columns (total 11 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   House Number                    6000 non-null   int64
 1   Street                          6000 non-null   object
 2   First Name                      6000 non-null   object
 3   Surname                         6000 non-null   object
 4   Age                             6000 non-null   object
 5   Relationship to Head of House   6000 non-null   object
 6   Marital Status                  4719 non-null   object
 7   Gender                          6000 non-null   object
 8   Occupation                      6000 non-null   object
 9   Infirmity                       6000 non-null   object
 10  Religion                        4681 non-null   object
dtypes: int64(1), object(10)
memory usage: 515.8+ KB
```

Heat map showing the null values among the datasets.

Religion, Dealing with Missing Values

The missing religion values for minors were replaced with their parent's religion since it's likely they practice the same religion. 'Private' religion choices were kept private as it's important to respect respondent privacy. The blanks were replaced with NA.

Cleaning Marital Status

I filled in missing marital status values for minors under 16 with 'Not Applicable', as they are not legally able to marry. For individuals over 16, I filled in missing marital status values with 'Missing' because, according to the Marriage Act of 1949, it is possible to marry with parental consent. Therefore, I used 'Missing' to indicate that the information is not available rather than assuming that they are unmarried.

Cleaning Relationship to head of house

Three households led by 16- or 17-year-olds with children were found, including a divorced 16-year-old with a child, a 17-year-old single parent, and a 16-year-old married with a child. These young individuals are likely responsible for their household and family members, and it's crucial to provide them with proper support

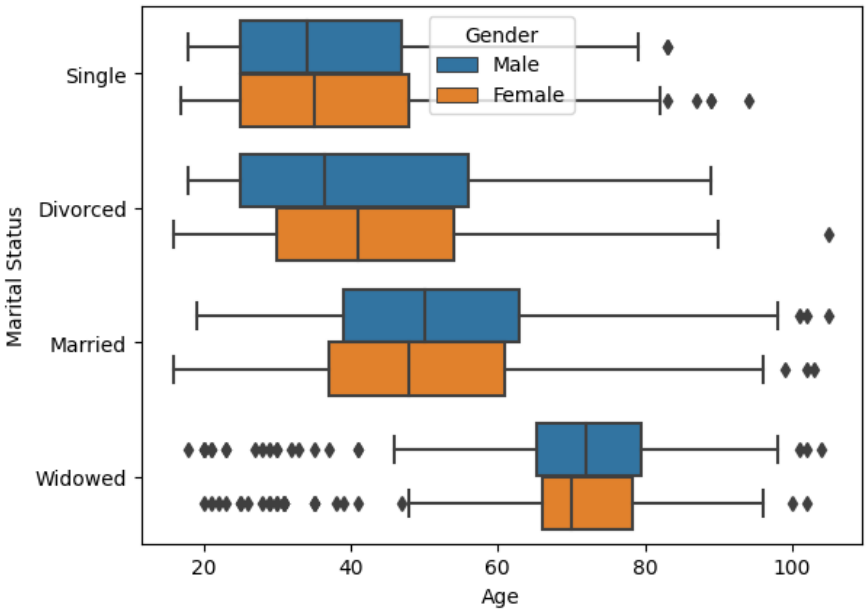
and resources such as education, healthcare, and legal services. A misspelling of "Niece" was corrected from "Neice".

Cleaning First Name, Infirmity, Occupation and others

Blank in ‘First Name’ was replaced with ‘Missing’ as it is not possible to infer like the surname. Blanks in ‘Infirmity’ were replaced with ‘NA’. The two blanks in the occupation column were replaced with ‘Missing’ as the occupation status of these individuals is unknown. The occupation categories were also subdivided into main categories of Employed, Retired, Unemployed, Student, University student, child.

Handling Outliers

An outlier is a data point that differs significantly from other observations. Outliers could be as result error during data gathering. Outliers were identified by using Z-score and boxplot.



By visualizing the outliers in the marital status by the ages and according to the genders, it identified widows at a very young age as outliers especially male and this also showed that some were married and divorced at the age of 18, this could be unusual but not necessarily inappropriate.

Identifying Outliers for Age

I used a Z-score threshold of 3 to identify outliers in the age column. Ages over 100, such as 102, 103, 104, and 105, were initially flagged as outliers, but upon further review of the occupation and relationship columns, these ages appeared to be accurate. With the increasing trend of longer life expectancy, it is not uncommon for people to live beyond 100 years old.

```
# using Z score to identify outliers in the Age Colunm
Mean=df['Age'].mean()
Std=df['Age'].std()
threshold = 3
outlier = []
for i in df['Age']:
    z = (i-Mean)/Std
    if z > threshold:
        outlier.append(i)
print('Outlier for the Ages are', outlier)
```

```
Outlier for the Ages are [102, 102, 102, 105, 105, 102, 103, 104]
```

BIRTH RATE

From the analysis, the current crude birth rate of a town is 8.67 births per thousand, which implies that 8.67 babies are born per year for every 1,000 people in the town. To calculate the total number of births in the town, we can divide the total number of births in one year by the town's population and then multiply the result by 1000.

$$\frac{\#\text{ births in 1 year}}{\#\text{ thousand total population}} = \text{Crude Birth Rate}$$

Furthermore, four years ago, the birth rate per 1000 was estimated to be 11.83 after adjusting for the population in the previous year. This indicates that in the past four years, there has been a decline of 3.16 births per thousand in the town's birth rate.

This decline in birth rate may be attributed to various factors, including women focusing more on their education and careers, medical advancements such as lower child mortality, and access to contraception. The data also suggests that the town's population growth has slowed down due to the declining birth rate.

In summary, the current birth rate of 8.67 births per thousand in the town implies that the population is not growing.

DEATH RATE

The death rate of a population can be calculated by dividing the total number of deaths in one year by the absolute population and then multiplying the result by 1000. The analysis shows that the death rate in the population is 16.5 per thousand, which indicates that 16.5 individuals per 1000 people in the population die every year.

$$\frac{\#\text{ deaths in 1 year}}{\#\text{ thousand total population}} = \text{Crude Death Rate}$$

Factors such as better healthcare, low infirmity rates, and increased home creation have contributed to the decline in death rates to 16.5. This suggests an aging population and a need to prioritize homes for them. Improved medical knowledge, income, and housing quality have also lowered mortality rates.

FERTILITY RATE

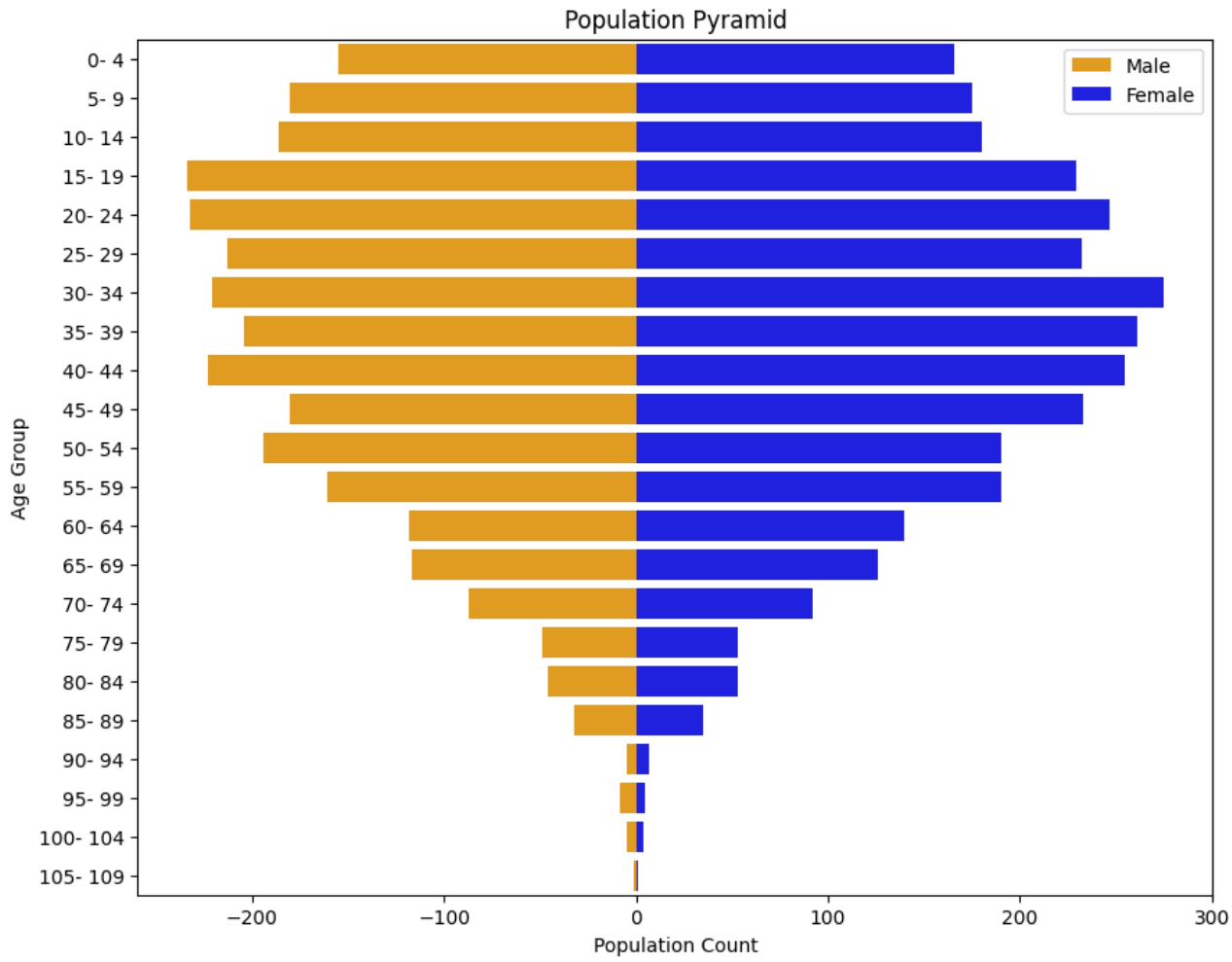
According to the analysis, the average number of children per woman of childbearing age in every 1000 is 226. This means that for every 1000 women of childbearing age, there is an average of 226 children born.

Total Fertility Rate (TFR)
= average number of births per woman per lifetime

Furthermore, the data suggests that women between the ages of 30-34 gave birth to more children than any other age group. This trend could indicate a delay in women giving birth, as women in this age group may be prioritizing other aspects of their lives, such as career or education, before starting a family.

It is essential to note that delaying childbirth can have implications for fertility and pregnancy outcomes. Women who delay childbirth may face challenges in conceiving or may have a higher risk of complications during pregnancy and childbirth.

POPULATION PYRAMID



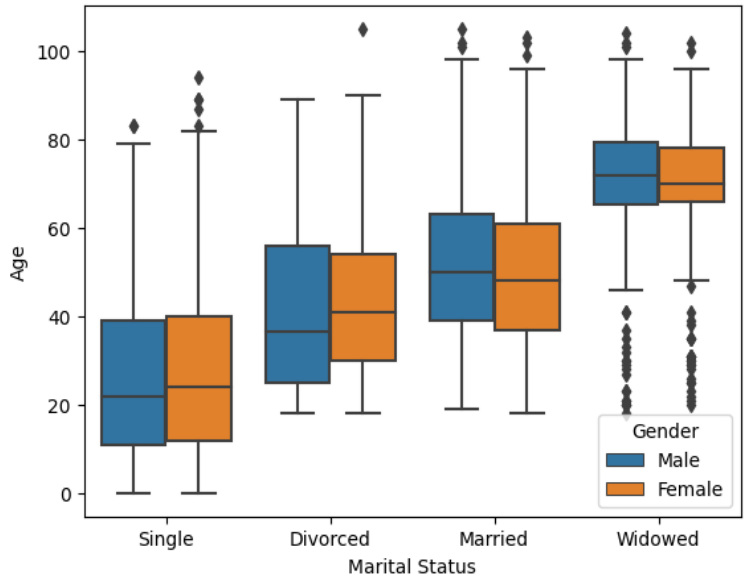
From the population pyramid of the city, it appears that there is a relatively low birth rate within the age range of 0-4 years, compared to the number of individuals within middle age brackets. Additionally, there seems to be an increasing number of individuals aged over 65 years, which is indicative of a low death rate within the city.

Given these observations, it is likely that the city is experiencing slow or no long-term natural increase. This is because the low birth rate implies that there are fewer

new births to increase the overall population, while the low death rate implies that individuals are living longer and contributing to the aging population. Thus, the population growth in the city may be primarily influenced by migration.

DIVORCE AND MARRIAGE RATE

From the analysis, the divorce to marriage ratio is 34.56 per 1684 marriages, indicating a relatively high divorce rate. Out of the total of 1684 marriages, there were 582 divorces recorded.



To calculate the divorce rate, the number of divorcees is divided by the total number of marriages (1684) and multiplied by 100, resulting in a divorce rate of 34.56%. This figure suggests that over one-third of marriages end in divorce.

It is also noteworthy that the number of female divorcees is higher than that of male divorcees, indicating that women are more likely to initiate divorce proceedings.

IMMIGRATION AND EMMIGRATION

Immigrants are individuals who move to a different country, and according to the data, there has been an increase in the age group of 15-59 years and a decrease in the age group of 25-29 years. This suggests that many students come to the town to attend university and then leave upon graduation. However, those who decide to stay may do so due to the availability of job opportunities and a low cost of living.

To calculate the immigration rate, the number of lodgers and visitors was used, resulting in 27 immigrants per thousand people in the town. This number highlights the attractiveness of the town as a destination for individuals seeking better opportunities.

On the other hand, emigration was calculated based on the difference between male and female divorcees. The data shows that there are more female divorcees than male divorcees, indicating that men are more likely to leave the town after a divorce. From the calculation, the emigration rate is estimated to be 18 per thousand people in the town.

In conclusion, the data suggests that the town attracts a significant number of immigrants, particularly students attending the nearby university, due to job opportunities and a low cost of living. However, the town also experiences a level of emigration, particularly among men after a divorce.

NATURAL RATE

Natural Increase Rate	=	Birth Rate	-	Death Rate

According to the data, the natural increase rate in the city is -7.83%, which is the difference between the death rate and birth rate. This indicates that the rate of birth is less than the rate of death within the year, resulting in a negative rate of natural increase. This negative trend suggests that the city is experiencing a decline in population growth.

NET MIGRATION

$$\text{Net Migration Rate} = \text{Immigration Rate} - \text{Emigration Rate}$$

According to the data, the net migration rate in the city is 8.17%, which is the difference between the immigration rate and the emigration rate. This indicates that more people are coming into the city than leaving, which can help to stabilize the negative natural increase rate (birth rate - death rate).

The positive net migration rate suggests that the city is attractive to immigrants, potentially due to factors such as employment opportunities, affordable housing, and access to education and healthcare. This influx of new residents can help to offset the negative impact of the declining population growth rate, as it contributes to an increase in the overall population.

POPULATION GROWTH

The growth rate of a population refers to the percentage change in population size over a specific time period. It is influenced by a variety of factors such as immigration, emigration, birth rate, and death rate.

$$\text{Population Growth Rate} = \text{Natural Increase Rate} + \text{Net Migration Rate}$$

Based on the available data, the total population of the city has grown by 0.34% over the specified time period. This growth rate can be largely attributed to the positive net migration rate, as the declining fertility rate has resulted in a decrease in the natural increase rate.

Commuters

The identification of commuters in the town was conducted using a methodology that consisted of two criteria: individuals who self-identify as University Students, including PhD students, and those who have occupations that require commuting. Occupations such as baristas, community roles, and retail workers were deemed non-commuting jobs.

Based on this approach, the study identified a total of 2256 commuters, comprising 37.6% of the entire population. This percentage is relatively high and may have implications for the transportation infrastructure and traffic patterns in the town.

Occupancy Rates

Based on the analysis conducted, the mean occupancy rate of the houses is approximately 3 (specifically 2.57). This indicates that, on average, there are about 3 individuals living in each house. However, further analysis reveals that there are approximately 959 houses that are over-occupied out of the total 2330 houses examined.

Moreover, there were cases where a single house was occupied by over 20 individuals. This is indicative of an affluent family, as it is expected that a family with such a large number of members would require a massive house to accommodate everyone comfortably.

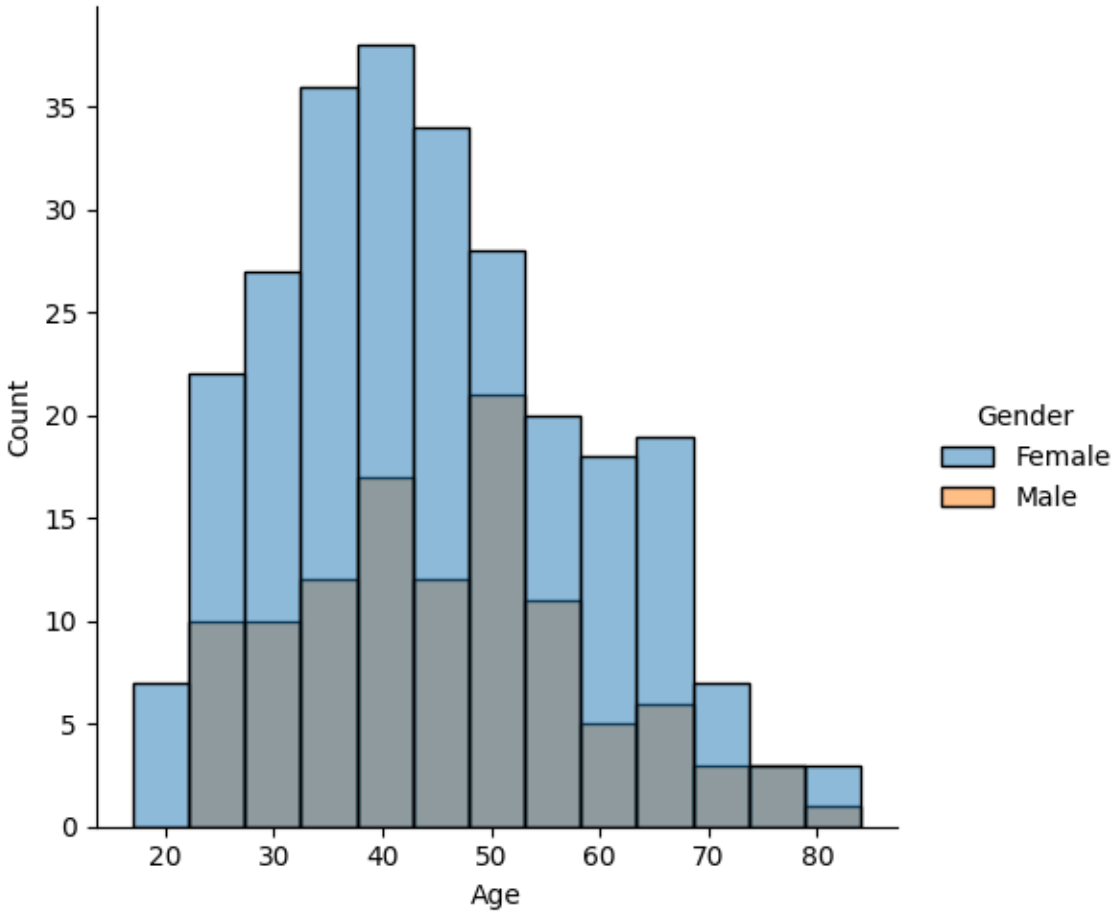
```
# occupancy describe
occupancy['No of Occupant'].describe()
```

```
count    2330.000000
mean        2.575107
std         1.652685
min         1.000000
25%         1.000000
50%         2.000000
75%         4.000000
max        20.000000
Name: No of Occupant, dtype: float64
```

Unemployment Rate

Based on the calculations, the unemployment rate in the city is 8.98%. This rate indicates the proportion of individuals who are currently not employed in relation to the total workforce. A high unemployment rate has several negative effects on the economy, including a reduction in real income for people and a decrease in the overall output of the economy.

$$\frac{\text{Unemployed People}}{\text{Labor Force}} \times 100$$



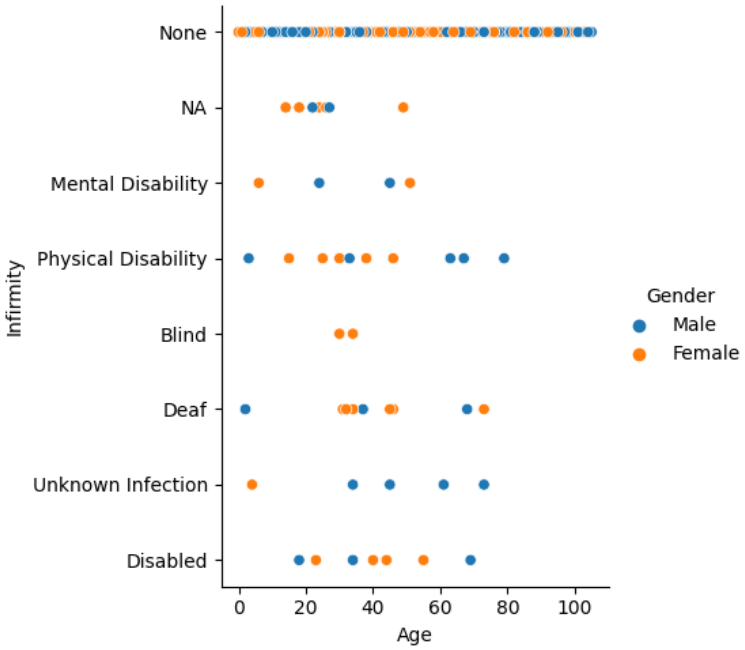
To address the high unemployment rate in the city, the government should focus on creating jobs and improving skills development for women and individuals aged 30-50. Access to education and training programs should also be improved.

INFIRMITY

The low infirmity rate of less than 1% in the town suggests that there are standard healthcare services available to the residents. However, upon closer inspection, it is apparent that there are still a few disabled people, especially children in the town who require additional support and care.

It is important for the government to take action to support the young disabled children and the old-aged people, ensuring they receive the care and resources they need to live their lives to the fullest. The percentage of infirmity, while relatively small, is a sensitive and important issue that should not be ignored.

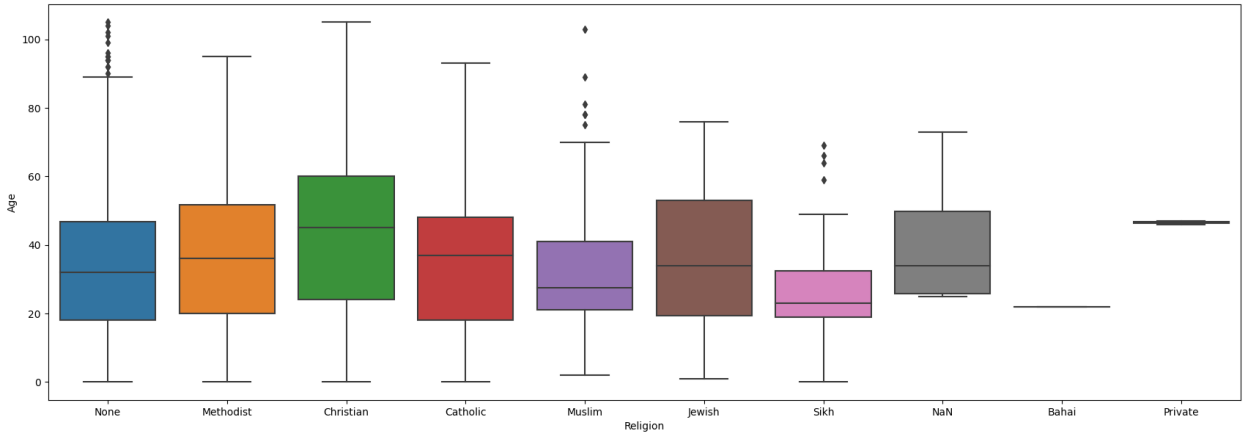
Value	Count	Frequency (%)
None	5960	99.4%
Physical Disability	10	0.2%
Deaf	9	0.2%
Disabled	7	0.1%
Unknown Infection	5	0.1%
Mental Disability	4	0.1%
Blind	2	< 0.1%



The plot shows the relationship between Infirmity and Age

Religion

From the analysis, it appears that the town is not a religious community, with a high number of respondents indicating 'None' as their religious affiliation. However, among those who provided a specific religious affiliation, Christian has the highest number of followers, followed by Catholic. It is also worth noting that Muslim appears to be an emerging religion in the town, which may be attributed to immigration.



Other religions currently have small followers in the town, so building a new church may not be a priority above other needs of the population.

Conclusion and Recommendation

To benefit both the commuting student population and the workforce, I recommend that a train station should be built with the occupied piece of land. This could attract investment and boost the town's economic growth. With an aging population, it is crucial to prioritize investment in age-old care services to meet the changing needs of the community and adequately prepare for future demand.

References

Anon., 2014. Marriage Act 1949. [Online]

Available at:

https://www.legislation.gov.uk/ukpga/Geo6/12-13-14/76/section/3

[Accessed 27 April 2023].

Office for National Statistics. (2021). Births in England and Wales: 2020. [Online]

Available at:

https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsummarytablesenglandandwales/2020

[Accessed 27 April 2023].

Department for Education. (2021). Characteristics of children in need: 2020 to 2021. [Online]

Available at:

https://www.gov.uk/government/statistics/characteristics-of-children-in-need-2020-to-2021

[Accessed 27 April 2023].

Eurostat. (2021). Fertility statistics [Online]

Available at:

https://ec.europa.eu/eurostat/statistics-explained/index.php/Fertility_statistics#Total_fertility_rate

[Accessed 27 April 2023].

Office for National Statistics. (2021). Deaths registered weekly in England and Wales [Online]

Available at:

https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregisteredweeklyinenglandandwalesprovisional/weekending22april2022

[Accessed 27 April 2023].

Office for National Statistics. (2021). Migration Statistics Quarterly Report: February 2021. [Online]

Available at:

https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/bulletins/migrationstatisticsquarterlyreport/february2021

[Accessed 27 April 2023].

The Economic Costs of Unemployment" from the International Labour Organization(2021)

[Online]

Available at;

https://www.bankofengland.co.uk/working-paper/2021/the-long-term-effects-of-unemployment

[Accessed 27 April 2023].