# CENSUS PROJECT REPORT

# THE PROJECT TASK-

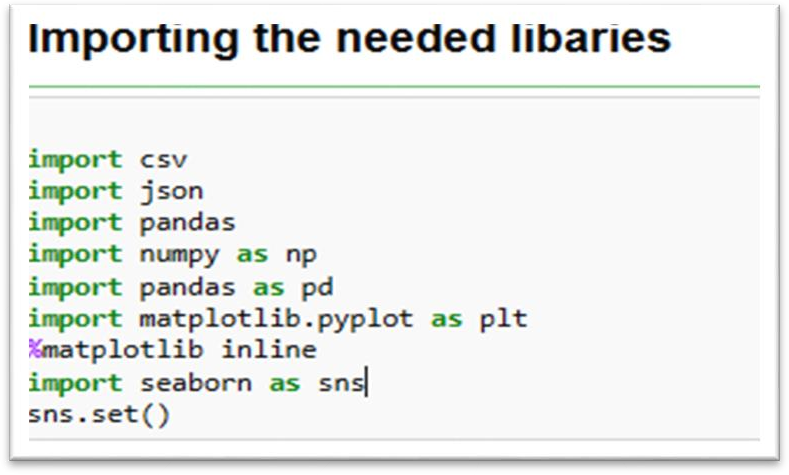
# The town from the census is a modestly sized one sandwiched between two much larger cities that it is connected to by motorways. The town does not have a university, but students do live in the town and commute to the nearby cities. Once you have a cleaned dataset to analyze, your task, you have to decide the following: (a) What should be built on an unoccupied plot of land that the local government wishes to develop?

# (b) Which one of the following options should be invested in? (i) Employment and training (ii) Old age care (iii) Increase spending for schooling

1**.0 INTRODUCTION**

This report aims to analyze mock census data06 for an average -sized town, with the goal of providing recommendation to the government about potential development on a vacant land parcel and future service investments to invest in the town. This process involves the scrutinizing and rectifying data errors, analyzing and presenting the data in a visual format, interpreting the findings and making informed conclusion and recommendations.

# 2.0 IMPORTATION OF LIBRARIES

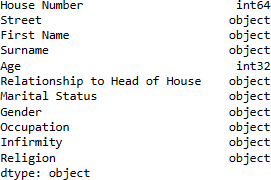
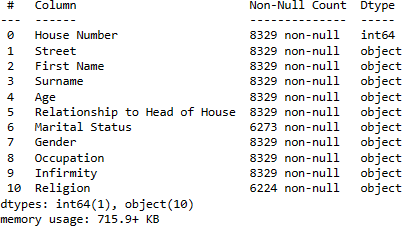


The needed libraries were imported to enable us to access varieties of arrays and functions and tools.

# DATA CLEANING

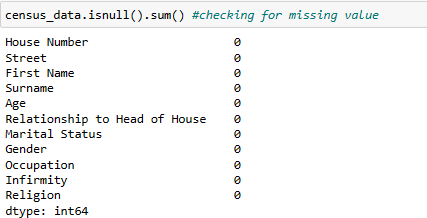
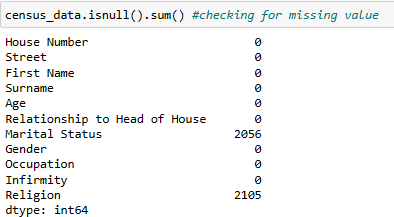
Absolute care was taken in the data cleaning process, which entailed the meticulous inspection of each data column. In the process of tidying up the data, series of checks were conducted, which includes:

* 1. **DATA TYPES**: Upon careful consideration of data type of each column, it was observed that age was needed to be converted to integers.



**Fig1**. Information about the data frame .

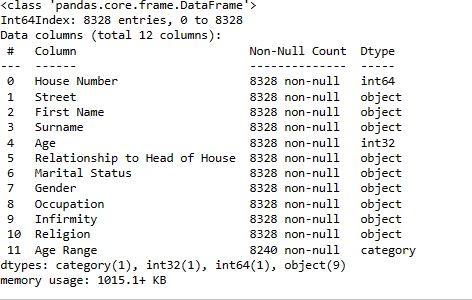
* 1. **Dropping entries:** one person was taken out of the dataset, as we cannot logically place a first name to that entry under the ‘’first name’’ column(variable).
  2. **Null and blank values:** The religion and the infirmity columns had the most null values. The religion’s variable missing values were changed to ‘’Unknown’’ and null values (NA) were changed to None, with exception to one entry(person) whose religion was changed to methodist as the family head was a methodist, so it was logical to assume he is of methodist doctrine. The infirmity had some missing values which I replaced with ‘unknown infection’, since we could not establish a relationship on what to replace it with. Similarly, the marital status three variable blank values were replaced with married, single and single respectively, as the first person was of 52years old, the head of the family and a practicing doctor, while the other two blank values were replaced with single as they were still 19 years and still students, as it was logical to do so. In addition, the marital status with null values whose ages fall below 18years were replaced/changed to ‘’below age’’ as by law (The Marriage and Civil Partnership (Minimum Age) Act 2022) they were not allowed to be married below 18years. Furthermore, the age variable(column) had two blank values, which replaced with 35 which is the mean age of the population and converted the ages from floats to integer.



**fig2.** Data frame with null values and Data frame without null values

In order to obtain a more accurate analysis of the number of retired and unemployed individuals, the government has changed the classification of individuals who are above the age of 65 and unemployed from ‘’unemployed’’ to ‘’ Retired’’ in their records, as the retirement age is set to 65 by the government. (Gov.uk).

The finalized census data after cleaning will have the following characteristics:



**Fig3** Finalized data frame

To help in this report analysis, the below terms were used:

* **Age range**: Individuals Age placed in a 5-year age bands for population pyramid.
* **Job Status:** Simplified occupations has the values: Student (Child), Employed, Unemployed, Student, Retired.
* **Home occupancy**: the number of all individuals in a household.

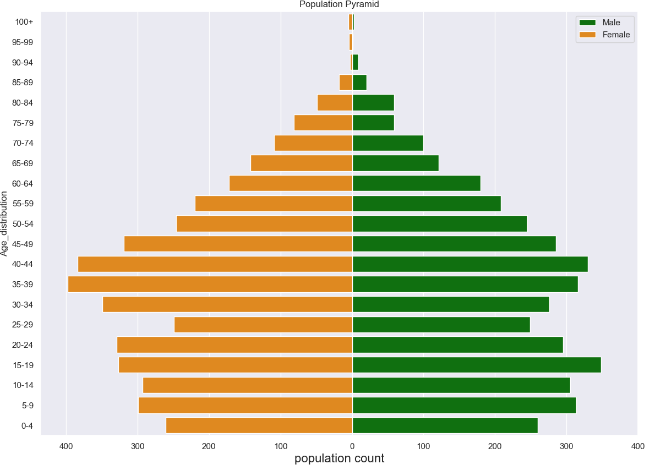
# EXPLORING DATA THROUGH DATA ANALYSIS AND VISUALIZATION:

* 1. **Gender**: Within the dataset, there is a higher proportion of females more than the male, with females accounting to 51.7% (4302), while the male comprised of 48.3% (4026).



**Fig4** Population count of the male and female in the population

* 1. **Age:** The age population pyramid is a visual depiction of the age distribution among male and female members of a population at various time points. This pyramid does not only display the percentage of younger individuals (such as infants and children), middle aged individuals and older individuals, but it can also project future population trends.



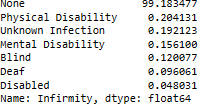
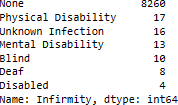
**fig5**. Population Pyramid

From the above population pyramid structure, the analysis indicates a relatively bigger proportion of young babies particularly those aged 0-4 when compared to the aged people, implying a gradual growth rate in the population. Moreover, both male and female populations have a tendency to live longer and attained old age.

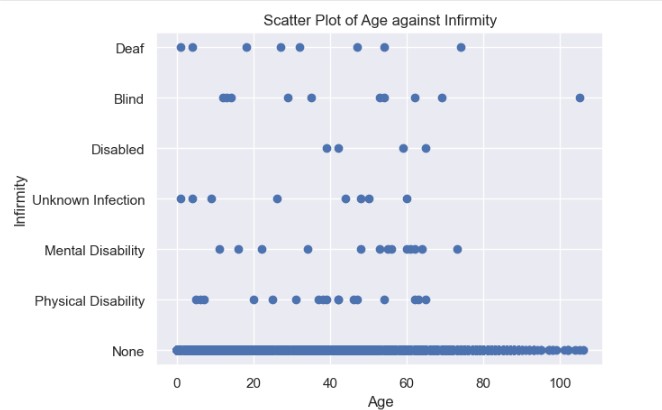
Upon conducting a more detailed descriptive examination, it becomes evident that the majority of the town’s inhabitant are employed accompanied by a significant number of school-aged children, while the proportion of unemployed individuals stands at approximately 5.67%.

Additionally, a large portion of the population identifies as either married or single, and in terms of infirmity, the prevalence is relatively low when compared to other factors; the proportion of individuals with infirmity is 0.72%, which is less than 1% of the total population.

# Infirmity:



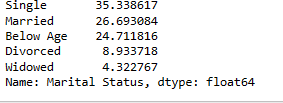
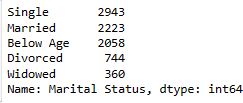
**Fig 6** Infirmity description and Infirmity distribution in percentage%



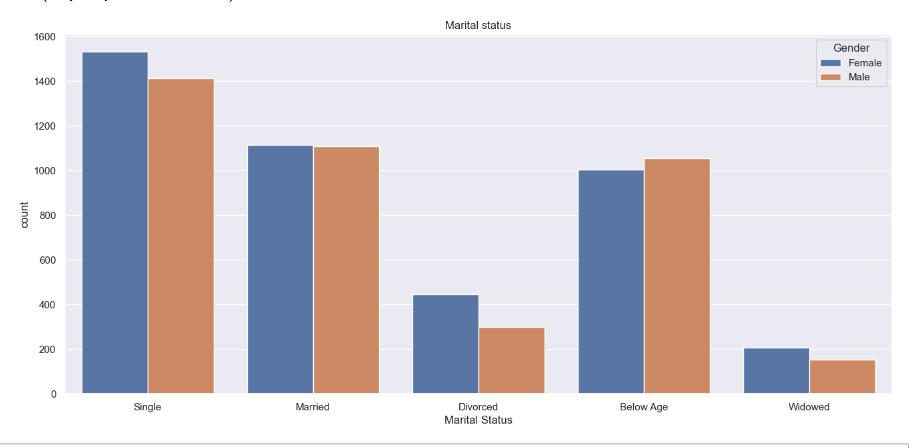
**Fig 7** scatter plot of age against infirmity

The fact that only 0.7% of the population fell ill or are physically challenged in one way or the other suggest that the town is generally healthy, with about 99.3% of the population being in good health. When the age range of those who became sick was examined, it was found that individuals over 80 years old of age were largely unaffected, with only one case report as shown in the scatter plot diagram above. This finding will be taken into account when deciding what type of medical facilities should be constructed in the city in the future.

# Marital Status :



**Fig8** Marital status description and Marital status distribution in percentage%

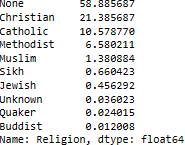
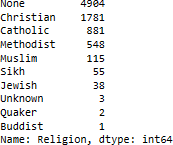


**Fig. 9**. Histogram for the count of marital status

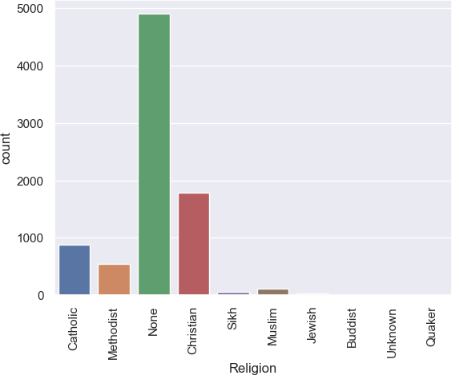
The largest proportion of the population, accounting for 35.3%, consists of individuals who are single, followed by married individuals who make up 26.7% of the population.

The data also reveals that in the city, the ratio of female divorcees to male divorcee is 445 to 299, with the female having more divorce rate than the male as shown in the graph above.

# Religion:



**Fig10.** Religion description and religion distribution in percentage%



**Fig. 11**. Histogram for the count of religion

From the above, it is observed that Christianity is the most practiced religion with percentage of 21.4% of the population and this is followed by the catholic and other faiths in descending order. A large proportion of the population identifies themselves as having no religion or holding a religious value of ‘’’None’’. It is crucial to consider these beliefs when making a decision about whether or not to construct an additional worship center.

# Birth rate and death rate:

The formular below were used to calculate the birth and death rate per thousand of the population .

𝐵𝑖𝑟𝑡ℎ 𝑟𝑎𝑡𝑒 = 𝑁𝑢𝑚𝑏𝑒𝑟 𝑜𝑓 𝐵𝑖𝑟𝑡ℎ 𝑥 1000

𝑇𝑜𝑡𝑎𝑙 𝑝𝑜𝑝𝑢𝑙𝑎𝑡𝑖𝑜𝑛

𝐷𝑒𝑎𝑡ℎ 𝑟𝑎𝑡𝑒 = 𝑁𝑢𝑚𝑏𝑒𝑟 𝑜𝑓 𝐷𝑒𝑎𝑡ℎ 𝑥 1000

𝑇𝑜𝑡𝑎𝑙 𝑝𝑜𝑝𝑢𝑙𝑎𝑡𝑖𝑜𝑛

The current birth rate is approximately 11births per thousand (1.1%), while the current death rate is approximately 1death per thousand (0.1%). The birth rate is gotten by counting the number of age that is equal to zero, dividing by the population length and multiplying by the 1000 per year population value used. The death rate involves first of all knowing the life expectancy of people in the UK , which is 82years as at 2023. I found the difference between two age bands(78-83) and84-89) which fell between the 82years life expectancy ,then divide by the population length used and multiplying by 1000 per year population value used and dividing the overall result by 5(Age band). The result obviously indicates that there are more births than deaths.

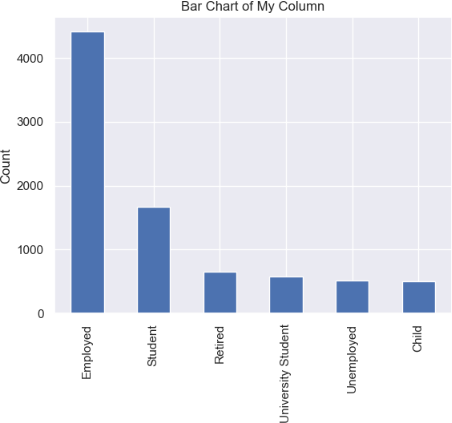
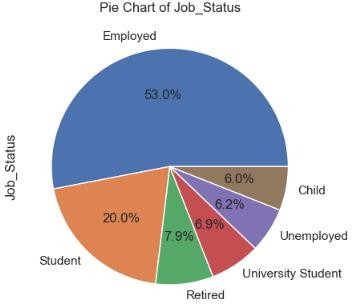
# 4.7.MIGRATION:

The majority of emigration and immigration in the town can be attributed to university students. Students can be observed to be a constant in the town’s growth, as their respective houses will always be occupied each year whenever they leave.

Immigration statistics are gotten from the visitors and lodgers who are single, with exception to those who are no longer married and are lodging after leaving their spouses. Emigration statistics are determined by analyzing the difference between female and male divorcees.

The emigration per thousand is approximately 18(2.11%) and the immigration per thousand is approximately 52(5.2%). This shows putting into consideration, the whole population, more people move into the city more than the number that are moving out of the city.

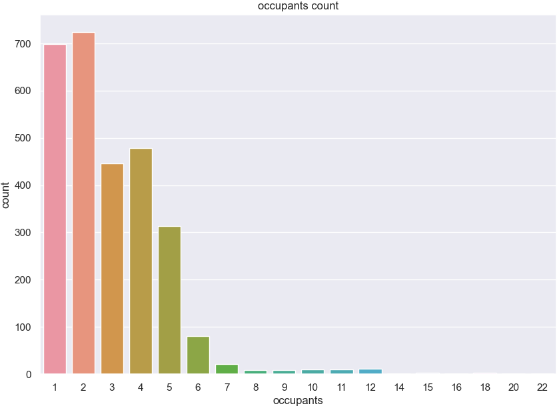
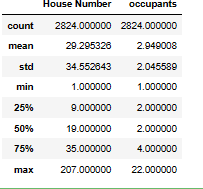
# EMPLOYMENT:



**Fig. 12**. Pie chart showing the percentage of occupation **Fig. 13**. Bar chart showing the count of occupation

From fig 12 and fig 13, it can be observed that Less than 8% of the population are unemployed, while 53% of the population are employed. 26% of the population consist as students. It is imperative to consider this employment statistics as it will play a role in determining whether the town needs to invest in employment and training or not.

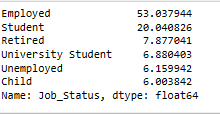
# OCCUPANCY RATE:



**Fig. 14**. Household ocupants description **Fig. 15**. Histogram for the count of ocupants

From fig 14 and fig 15 respectively, we can see that the average number of occupants living in a household is approximately 3persons per household as shown by the mean, which shows that the house is not crowded. It is important to consider these household number of occupants as it will be vital when making a decision about whether or not to construct a high- or low-density house building.

# COMMUTERS:



**Fig. 16**. Dataframe showing percentage distribution of students and some of the working class

As observed in fig 16 , it is estimated that there are estimated 27% of the population are students while Then larger part of the population are working class constituting around 53% of the total population. Commuters logically were identified as university students and most of the employed population as they are the teeming part of the population that will be commuting between towns and outside of the town for their various jobs and school.

# 5.0 RESULT, DISCUSSION AND RECOMMENDATION:

Approximately 32% of the population identified themselves as Catholic and Christians, and the town already has a church that serves Catholics. Thus, there isn't a pressing demand to build another church.

Looking at our infirmity data of **fig6**, over 99.3% of the population now enjoys good health, hence no dire need of any emergency medical building.

There is now no need to raise the amount of money spent on education or make investments in fundamental infrastructure because the town is not significantly expanding.

According to data, there is an average of 3persons per household in the town and the populations natural rate of growth is not massively growing. The percentage increment is actually not huge and an average household in the town is not crowded, hence, there isn't a pressing need to build high or low-density housing.

In addition, larger proportion (over 53%) of the population are employed, and less than 8% are unemployed, this shows that investment in employment and training might not be the most needed sector to fund in at the moment as the town is currently not hit with high unemployment.

# 6.0 CONCLUSION:

1. After careful analysis and discussions, my conclusion is that the vacant plot of land should be used for the construction of train station as there is no train station in the town. It is obvious that there is high density of working-class population with percentage of 53% as shown in fig 12 and approximately 27% of the population who are students. This will ease the everyday commuting of different employed people who are working in different sectors within the town and the neighboring towns and the students who will be needing it too to go to their everyday schools.
2. It is also evident from the population pyramid(**fig5**) that there will be increased number of retired people in the nearby future, there will be need for age care home in the future years, and as such the town needs to start working towards allocating more funds to end of life care.

# 7.0 REFERENCES :

*Eurostat (2017) Marriage and Divorce Statistics*

Available online: <https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Marriage_and_divorce_statistics>

[Accessed 26/04/2023].

*Gov.uk (2023) Implementation of the marriage and civil partnership(minimum age)Act 2022*

Available Online: [https://www.gov.uk/government/news/implementation-of-the-marriage-and-civil-partnership-minimum-age-](https://www.gov.uk/government/news/implementation-of-the-marriage-and-civil-partnership-minimum-age-act-2022) [act-2022](https://www.gov.uk/government/news/implementation-of-the-marriage-and-civil-partnership-minimum-age-act-2022) [Accessed 10/04/2023].

*U.K life Expectancy(1950-2023)*

Available online: [https://www.macrotrends.net/countries/GBR/united-kingdom/life-](https://www.macrotrends.net/countries/GBR/united-kingdom/life-expectancy#%3A~%3Atext%3DThe%20current%20life%20expectancy%20for%2Ca%200.15%25%20increase%20from%202020)

[expectancy#:~:text=The%20current%20life%20expectancy%20for,a%200.15%25%20increase%20from%202020](https://www.macrotrends.net/countries/GBR/united-kingdom/life-expectancy#%3A~%3Atext%3DThe%20current%20life%20expectancy%20for%2Ca%200.15%25%20increase%20from%202020) [Accessed 10/04/2023]

*UK Labour Market Statistics(2023)*

Available online : [https://commonslibrary.parliament.uk/research-briefings/cbp-](https://commonslibrary.parliament.uk/research-briefings/cbp-9366/#%3A~%3Atext%3DIn%20December%202022%20%E2%80%93%20February%202023%2Cbelow%20their%20pre%2Dpandemic%20level)

[9366/#:~:text=In%20December%202022%20%E2%80%93%20February%202023,below%20their%20pre%2Dpande](https://commonslibrary.parliament.uk/research-briefings/cbp-9366/#%3A~%3Atext%3DIn%20December%202022%20%E2%80%93%20February%202023%2Cbelow%20their%20pre%2Dpandemic%20level) [mic%20level](https://commonslibrary.parliament.uk/research-briefings/cbp-9366/#%3A~%3Atext%3DIn%20December%202022%20%E2%80%93%20February%202023%2Cbelow%20their%20pre%2Dpandemic%20level) [Accessed 01/04/2023].

*Retirement age in the UK(2023)*

Available online : [https://blog.moneyfarm.com/en/pensions/retirement-age-in-the-uk-when-can-you-retire-and-get-](https://blog.moneyfarm.com/en/pensions/retirement-age-in-the-uk-when-can-you-retire-and-get-your-state-pension) [your-state-pension](https://blog.moneyfarm.com/en/pensions/retirement-age-in-the-uk-when-can-you-retire-and-get-your-state-pension) [Accessed 25/03/2023].

*Birthrate vs death rate(no date) Our world in Data.*

Available online : <https://ourworldindata.org/grapher/birth-rate-vs-death-rate>[Accessed 29/03/2023].