UK CENSUS PROJECT REPORT

Table of Contents

1. INTRODUCTION 2

2. DATA CLEANING 2

3. POPULATION DEMOGRAPHICS 3

4. RECOMMENDATION 10

5. REFERENCES 11

1. INTRODUCTION

The United Kingdom (UK) carry out population census every 10 years. This is an important measure to control and ensure the country's growth and development. Policymakers have been unable to make proper land-use plans, develop policies, and allocate certain funding and investments due to a lack of knowledge about the population and distribution of residents in communities. The UK Federal Government distributed a land to this modest town to develop the town and improve urbanization. Some factors like household heads can sometimes provide false information thus compromising the data in the census. This can be corrected through efficient data analysis to detect and correct this false information according to other related data provided in the census.

The report includes cleaning up missing data and ensuring all our values are valid. In addition, some crucial analysis will be required to improve the quality of the given dataset. Also, visualizing the data using various plots and representations such as histograms, pyramids, boxplots, and scatterplots will be used to provide recommendations. Observations drawn from these plots will infer on what is best for the development of the proposed unoccupied land. However, the purpose of the census analysis is to recommend what kind of development should be made on the unoccupied plot of land, and what investment can be made.

2. DATA CLEANING

2.1 Data Features

The required python library was imported on the jupyter notebook. After importing the dataset, it was found out that a “Unnamed” column was unnecessary because it is just a duplicate of the serial number which will contribute nothing to the analysis. The census data consists of 11 features are shown in figure 1 below:

```
:  1 census_data.info() # checks for the entries, columns, nu
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10593 entries, 0 to 10592
Data columns (total 12 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   Unnamed: 0                    10593 non-null  int64
 1   House Number                  10593 non-null  object
 2   Street                        10593 non-null  object
 3   First Name                    10593 non-null  object
 4   Surname                       10593 non-null  object
 5   Age                           10593 non-null  float64
 6   Relationship to Head of House 10593 non-null  object
 7   Marital Status                7986 non-null   object
 8   Gender                        10593 non-null  object
 9   Occupation                    10593 non-null  object
 10  Infirmity                     10593 non-null  object
 11  Religion                      7911 non-null   object
dtypes: float64(1), int64(1), object(10)
memory usage: 579.4+ KB
```

Figure 1: The census data information showing the entries

The datatype of the age data series was given as a 'float' which is considered inappropriate because age is meant to be an integer. This was converted to 'int' function. From the earlier figure, the number of entries is incomplete for marital status and religion. These missing entries were fixed later in the cleaning process. In the case of house numbers, no blank entries were found but the word "Five" was changed to "5". This was to ensure that all its unique values are in figures not words. The only issue was with the 'First Name' column which was the space, the replace function was used to fix it appropriately.

Blank entries were present in surname, infirmity and religion. To fix this, the house address (street and house number) was checked to determine relationship to head and their household. In infirmity, blank spaces were assumed to be "None". People under age 18 were considered to have a religion so they replaced with "N/A". Some religion fields such as NaN, Nope, housekeeper, Quaker, Agnostic were considered as no religion (None). Verifying the unique values of Relationship to Head of House, "Neice" was wrongly spelt, it was replaced with "Niece" to avoid errors when indexing entries.

The minimum age of marriage is 18 years to protect children below that age from the scourge of forced marriage (Marriage Act, 2022). For this reason, marital status of underaged individuals was replaced by "NA" (underage). Errors ("nan", D, S, M, W) were replaced with the appropriate marital status entries such as Married, Divorced, Single, and Widowed. The house address was used to verify and replace underage people who were regarded as Head (Relationship to Head of House). The official full-time working age was set to 18 years because of vulnerabilities and constraints. However, people that fall between the ages of 16 and 17 are "Students" and could work part-time within the town. Employment status and age group columns were added to the census data set. People aged more than 65 are inspected and regarded as "Retired". The age group was used to develop an age population pyramid with the two genders (Male and Female).

Comments and markdown were used to show how the data cleaning and preparation for the detailed analysis of this report.

3. POPULATION DEMOGRAPHICS

By investigating the descriptive analysis of our data demographics, which involves data visualization to view the patterns and relationships among our features, the data gotten during the project were studied. Many people in the town are single or married, and the infirmity rate is very low when compared to other characteristics; so, it is deemed to be negligible.

3.1 AGE DISTRIBUTION

In this stage, a population pyramid is used which displays the age and gender distribution of the proposed town's population. This offers information regarding economic development and investment in the town (Indexmundi, 2017). Males (in blue colour) are positioned on the left and females (in red colour) are positioned on the right as the population is spread down the

horizontal axis. Based on trends in global migration, mortality, and fertility, the demographic pyramid gradually changes over time.

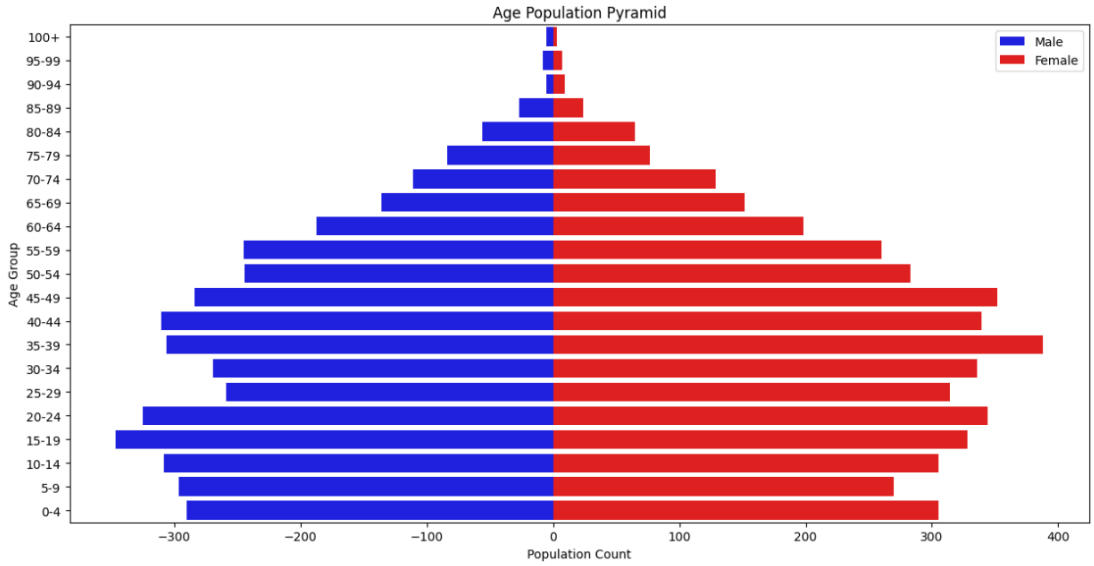


Figure 2: Age pyramid showing the population trends with respective age group

From the figure above, it is evident that age groups of (15 – 19) years and (35 – 39) years are the highest age and gender distribution, while people in groups of 100+ years are the lowest. This could also mean that there are slightly more middle-aged people than young ones because ages 0 to 14 are lesser.

3.2 EMPLOYMENT

The required population of the labour force are the people aged between 18 and 64 years which are University students, employed, and unemployed individuals.

University students are employed with other employed professionals. However, they amount to 53.2 % and 87.4 % of the town's total population and the workforce respectively, which means that they are likely to commute in and out of the town. People aged under 18 are considered as child and student considering individuals aged 65 and more to be retired.

Figure 3: Chart showing the employment status of different ages

Unemployment Rate

$$\text{Unemployment Rate} = \frac{\text{Unemployed Workers}}{\text{Total Labor Force}} \times 100\%$$

The descriptive analysis reveals that most of the town, of which 6 % of the census population are unemployed while the high unemployment rate is 9.8 % compared to the UK rate of 3 – 5 % (ONS, 2021), with a high percentage of university students and employed workers.

There is a relatively high unemployment rate in females than male (especially middle aged). Hence, the need to prioritize female as underrepresented gender and provide trainings and new skills for them as well as the males.

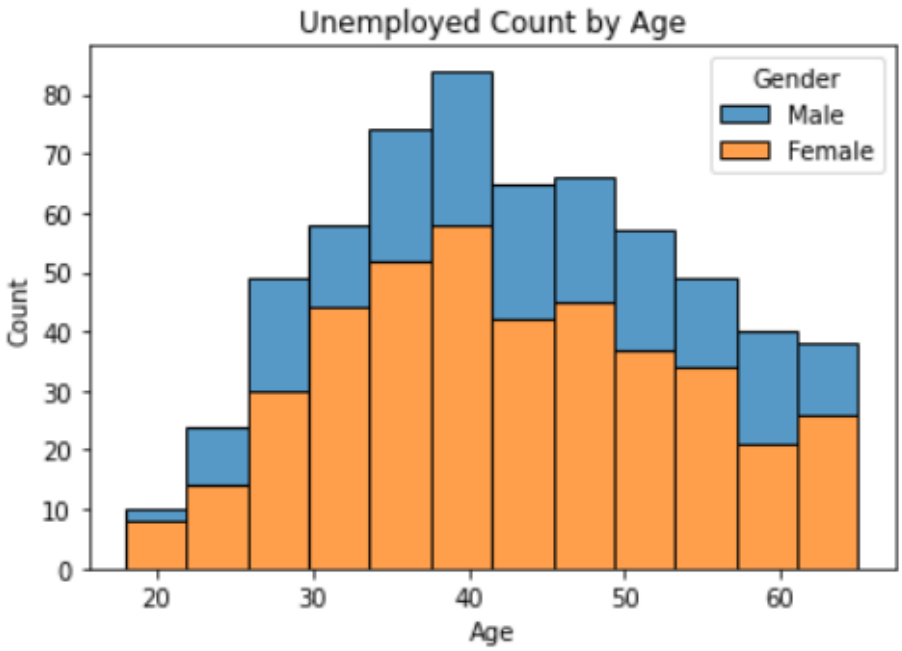


Figure 4: The distribution of unemployed people against age and gender

3.3. COMMUTERS

Commuters are likely to be Employed individuals, University students, professors and other professions that must travel to carry out their work. Students, Children, unemployed individuals, and retired people were not considered to be commuters as they do not have the need to travel in or out of the town. One of the most reliable modes of transportation in the UK is public transportation (Statista, 2022). Since salaries data was not given, it's highly recommended that buses and railway stations should be built considering there are potential 6,369 commuters.

Marriage and Divorce could affect the rate of potential commuters. There were 141 marriage rates per thousand while 94 divorce rates per thousand. The distinct between the two were not much. Divorce rate to marriage rate of 2 to 3 shows that there is high number of commuters because divorcees tend to travel more often than married individuals (subject to occupants and young dependants). Hence, this analysis contributes to further evaluation on potential commuters in the coming years.

3.4. OCCUPANCY LEVEL

This involves the number of people per household. There are 3,541 houses in the town, of which some houses are over-occupied. As shown in the figure below, there is a median occupancy rate of 3 which is higher than the UK median rate of 2.4 (ONS, 2021). It is not ideal because about 150

houses are said to be over-occupied. Three houses with 21 occupants each could be mistaken for an hotel or Airbnb. Regardless of this condition, low density housing is recommended or more motels could be built to balance the occupancy.

	Occupant(s)
count	3541.000000
mean	2.991528
std	1.879397
min	1.000000
25%	2.000000
50%	3.000000
75%	4.000000
max	21.000000

	No. of Houses
2	837
1	783
4	693
3	650
5	397
6	99
7	23
8	12
9	10
10	9
11	9
12	6
19	4
21	3
13	2
14	2
16	1
18	1

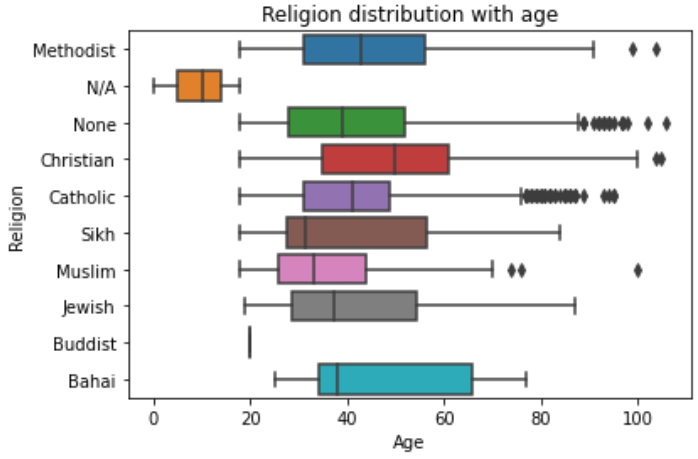
Figure 5: Table describing occupants per house

3.5. RELIGION AFFILIATIONS

The chart shows that people aged 18 years and below don't have a religion as they are minors (N/A). "Over time, there has been a dramatic decline in the proportion of people who identify with Christianity along with a substantial increase in those with no religious affiliation", which are classified as "None" (Curtice, J. et al, 2019). They amount to about 32 % of the town's population.

However, it is evident that about 23 % of the population are Christians (most dominant religion) with 51 % of them aged below 50 years.

None	3402
N/A	2669
Christian	2472
Catholic	1162
Methodist	637
Muslim	144
Sikh	64
Jewish	36
Bahai	6
Buddist	1



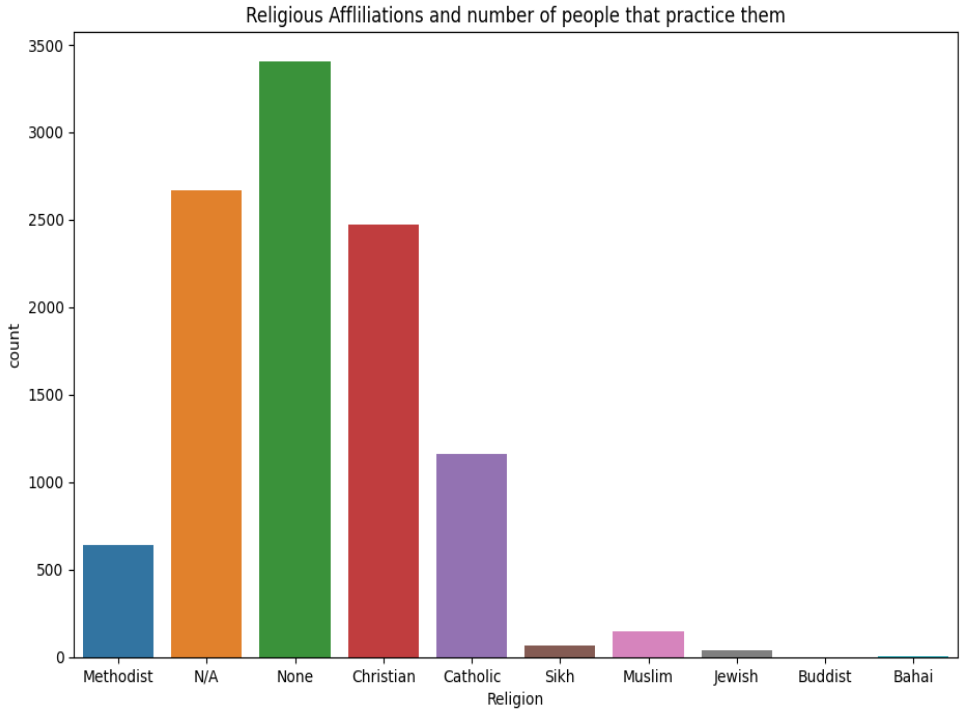


Figure 6: Charts showing number of people in each religions and ages

The Catholic and Methodist are the second and third most dominant denominations respectively with most people (that practice them) aged 35 and over. Although, people in other religion like Sikh, Muslim and Jewish are few but could increase in the coming years.

3.6. POPULATION GROWTH

Projections were based on the most recent mid-year population estimates for each UK country, as well as the most recent births, deaths, and migration data (ONS, 2021). This projection assumptions are based on past, observable long-term demographic patterns (ONS, 2021). This was derived from evaluating the town's population between four years ago and now. The natural

increase (birth and death rates) and net migration (immigration and emigration) were considered.

3.6.1 Birth Rate

$$\frac{births}{population} \times 1000 = birth\ rate$$

To determine the number of births, the people of age 0 were used. The expression above was used to evaluate the current birth rate of 10 births per thousand while the birth rate of four years ago is 11 births per thousand. This shows that the birth rate has been declining for the past four years. This was used to predict a declining birth rate in the next couple of years. However, other factors such as migration and fertility rate could reverse the situation.

3.6.1. Death Rate

$$\frac{deaths}{population} \times 1000 = death\ rate$$

"Across the UK, life expectancy at birth in recent years was a median of about 82 years" (ONS, 2021). To determine the death rate, individuals aged over 80 were used as number of possible deaths. The death rate is evaluated to be as 20 deaths per thousand while as at four years ago, it was 32 per thousand considering discrepancies. This decline predicts the declining death rate in the next few years because people over 65 years are likely to increasing life expectancy.

$$\frac{birth\ rate - death\ rate}{10} = natural\ increase\ rate$$

The natural population growth rate is declining by 1 because there is higher death rate than birth rate.

3.6.3. Migration

Immigration and emigration were determined using the number of university students, visitors and lodgers. Particularly, Single visitors and lodgers constitute of the proposed immigrants while emigrants are the single university students and divorced people (which are, of course, aged 18 and over).

$$\frac{\text{(\# of Immigrants - \# of Emigrants)}}{\text{Population/1,000}}$$

The rate of immigration constitutes of 38.3 per thousand while emigration rate amounts to 25.6 per thousand of the total population considering that people aged over 65 were excluded.

Evaluating using the expression (previous figure above), the net migration rate was analysed to be 12.7 per thousand.

$$\left(\begin{array}{c}\text{Births - Deaths}\\ \text{or } \mathbf{\text{Natural Increase}}\end{array}\right) + \left(\begin{array}{c}\text{Immigration - Emigrants}\\ \text{or } \mathbf{\text{Net Migration}}\end{array}\right) = \begin{array}{c}\mathbf{\text{Growth}}\\ \text{or decrease}\end{array}$$

Summing up the natural increase rate and net migration, it was concluded that there is a slight increment in population growth (of 0.25 %).

4. RECOMMENDATION

4. 1. DEVELOPMENT

Although, there is need for low-density housing, only few houses are over-occupied. Lodgers and visitors largely contributed to this and building houses might not improve the town's development. They could be potential immigrants and emigrants as well. Presently, there is no need to build religious building as the modal and average age of people practicing Christianity (most dominant denomination) are 56 and 49 respectively. Hence, there are fewer people aged below 50 who practice the mentioned religion which won't necessarily develop the town considering the younger age groups have no religion or practice other religions with fewer number of people. 99.4 % of the population have no form of infirmity, and the birthing age groups (18 – 45) are fewer as a result of decreasing birth rate. However, likelihood of emergency medical facility is not a necessity, now or in the future.

About 60% of the entire population are potential commuters being university students and employed professionals Also, single and divorcees (aged between 18 and 65) are commuters due to high divorce rate as stated earlier in the report. Travelling by rail has grown even in the face of rising passenger costs (GOS, 2019). It is highly recommended to build train station from the analysis.

4.2. INVESTMENT

Considering the employed people are largely more than unemployed people, only 6% of the whole population are unemployed. Providing training to unemployed individuals won't be recommended as an important investment as it is negligible in this case. Younger people from 0 to 15 years (school aged children) are few and as a result of declining birth rate and increasing death rate, investing in building primary schools is not recommended. Retired individuals (aged over 65) are about 10% of the total population and there are more odds of increasing number of older individuals in the next five years. However, investing in life-time old age care should be highly prioritised. Due to more commuters, there should be provisions of maintenance facilities for transportation and waste collection methods.

5. REFERENCES

GOS (2019). *A time of unprecedented change in the transport system: The Future of Mobility*. [online] Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/780868/future_of_mobility_final.pdf [Accessed 3 Dec. 2022].

Curtice, J., Clery, E., Perry, J., Phillips M. and Rahim, N. (eds.) (2019). British Social Attitudes: *The 36th Report, London.* The National Centre for Social Research. ISBN: 978-1-5272-4448-1, p2

Indexmundi.com. (2017). *United Kingdom Age structure - Demographics*. [online] Available at: https://www.indexmundi.com/united_kingdom/age_structure.html.

UK Legislation (2022). *Marriage And Civil Partnership (Minimum Age) Act 2022*. [online] www.legislation.gov.uk. Available at: https://www.legislation.gov.uk/ukpga/2022/28/notes/division/6/index.htm.

ONS (2021a). *National life tables – life expectancy in the UK - Office for National Statistics*. [online] www.ons.gov.uk. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2018to2020.

ONS (2021b). *Overview of the UK population - Office for National Statistics*. [online] www.ons.gov.uk. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/january2021.

Statista (2022a). *Average household size UK 2020*. [online] Statista. Available at: https://www.statista.com/statistics/295551/average-household-size-in-the-uk/.

Statista (2022b). *Modes of transportation for commuting in the UK 2021*. [online] Statista. Available at: https://www.statista.com/forecasts/997849/modes-of-transportation-for-commuting-in-the-uk.