TABLE OF CONTENTS

Census Project Report

Background

Every ten years, the United Kingdom undertakes a census of the population, with the most recent one having been conducted in 2021. The purpose of the census is to compare different people across the nation, to provide the government with accurate statistics of the population to enable better planning, to develop policies, and to allocate appropriate funding.

Objectives

(a) What should be built on an unoccupied plot of land that the local government wishes to develop? Choices are:

(i) High-density housing: This should be built if the population is significantly expanding.

(ii) Low-density housing: This should be built if the population is "affluent" and there is demand for large family housing.

(iii) Train station: There are potentially a lot of commuters in the town and building a train station could take pressure off the roads. But how will you identify commuters?

(iv) Religious building: There is already one place of worship for Catholics in the town. Is there demand for a second Church (if so, which denomination?), or for a different religious building?

(v) Emergency medical building. Not a full hospital, but a minor injuries center. This should be built if there are many injuries or future pregnancies likely in the population.

(vi) Something else?

(b) Which one of the following options should be invested in?

(i) Employment and training: If there is evidence for a lot of unemployment, we should re-train people for new skills.

(ii) Old age care: If there is evidence for increasing numbers of retired people in future years, the town will need to allocate more funding for end of life care.

(iii) Increase spending for schooling: If there is evidence of a growing population of school-aged children (new births, or families moving in to the town), then schooling spend should increase.

(iv) General infrastructure: If the town is expanding, then services (waste collection; road maintenance, etc.) will require more investment.

Data Cleaning

The census data had a lot of errors, missing values and a lot of anomalies, the data was cleaned using Jupyter notebook as a log book. Empty data were filled by assuming information from different households (e.g. Relationship to head of house), from different individual ('Gender') or ('Surname'). Some religions were replaced example of such is Jedi, Jedi isn't as a religion but a joke and it is encouraged to put 'No Religion' ("HumanistUk, 2021") thus, Jedi was replaced with 'None'. A larger percentage of NaN categories are minors, so they were replaced with Underage, we can conclude the minors are yet to decide which suits their belief. CAtholic, Meth-odist weren't spelt correctly so they were replaced with the correct spelling. Catholics, Methodists, Quaker, Baptist, Orthodoxy were replaced to Christians as they are been considered as different denominations under Christianity.

According to (NSPC, 2021) once a young person reaches 16 they can leave home or their parents can ask them to move out. However, parents are responsible for their children's wellbeing until they turn 18 - and they'll likely need support. There are four 'head of house' within the 16 – 17 years, it is assumed they have their parent's consent as they are married and they live together as seen on household relationship however they were replaced as lodgers because they are university students. Also, quite a number of NaN categories are University Student so they were replaced with Lodgers.

For Marital Status, over 2000 minors are under NaN category, they were replaced with 'Single', according to [Census, 2021] For most children under the age of 16, the response will be "Single (never married and never in a civil partnership) although according to the Ages of Marriage act of UK Parliament it is legal to get married at the age of 16 so far you get your parent's consent (Marriage Act, 1929).

There are two missing data for Age as they weren't inputted properly, a further check on the household couldn't give us the exact age but they are employed, female and are single so they were both replaced with the mean age of the female single in the community. There is an outlier of 120 years old, age of a

daughter whose dad was 36 years old; a further check on the household shows that they are twins and we can conclude there was an error of '0' at the back '12' as the other daughter is also 12 years old therefore this was replaced. There are a lot of widow 18 and 45 years, a simple code shows there are some widows that fall into 20 and below, we can conclude there are outliers and were replaced to single:
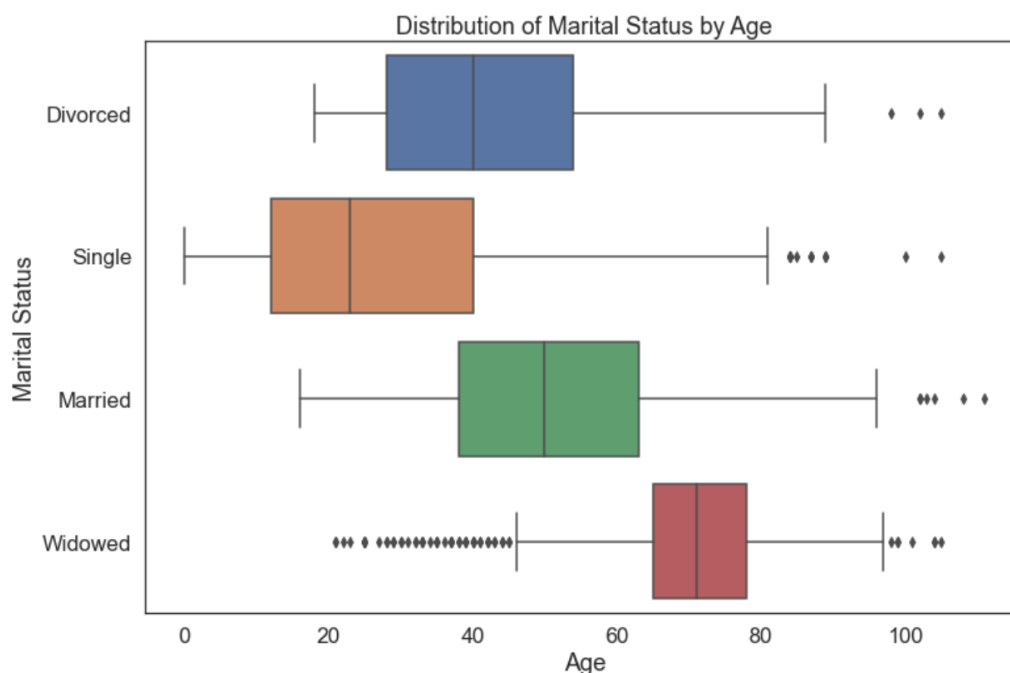
The graph shows unemployed count by Age, there are some ages above 65 that are unemployed, we can't categorically say that are unemployed as most might be retired; although one can keep working after one reaches State Pension age. 'Default retirement age' (a forced retirement age of 65) no longer exists ['(https://www.gov.uk/state-pension-age)', 2021]. There is also age for students so we can't conclude they are unemployed yet.

Unemployed Count by Age
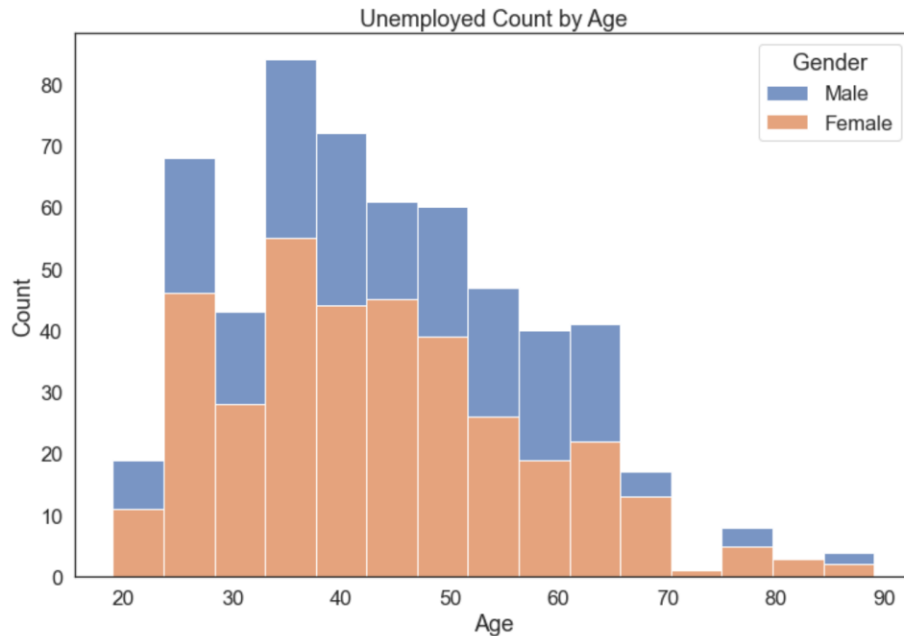
Demographics Data

Below shows the demographics data after cleaning the data which shows the data was thoroughly cleaned and missing data was filled appropriately:

```
House Number                    0        House Number              9099        House Number                    int32
Street                          0        Street                    9099        Street                         object
First Name                      0        First Name                9099        First Name                     object
Surname                         0        Surname                   9099        Surname                        object
Age                             0        Age                       9099        Age                             int32
Relationship to Head of House   0        Relationship to Head of House 9099    Relationship to Head of House  object
Marital Status                  0        Marital Status            9099        Marital Status                 object
Gender                          0        Gender                    9099        Gender                         object
Occupation                      0        Occupation                9099        Occupation                     object
Infirmity                       0        Infirmity                 9099        Infirmity                      object
Religion                        0        Religion                  9099        Religion                       object
Age Class                       0        Age Class                 9099        Age Class                    category
Employed                        0        Employed                  9099        Employed                       object
Occupation_Bar                  0        Occupation_Bar            9099        Occupation_Bar                 object
dtype: int64                             dtype: int64                          dtype: object
```

The below columns were added to carry out the analysis:

Occupation Bar: Simplified occupations with the values: Employed, Child, Student, University Student, Unemployed, Retired, PhD Student.

Age Class: Ages classed into 10-year age class.

The below Age Pyramid shows that there are lower number of people aged 9 – 0 in the young age group compared to the middle-aged group, this suggests a low birthrate. Furthermore, population of people in the elderly age group is low for both sexes which suggests that not a lot of people live into old age.

Age Pyramid
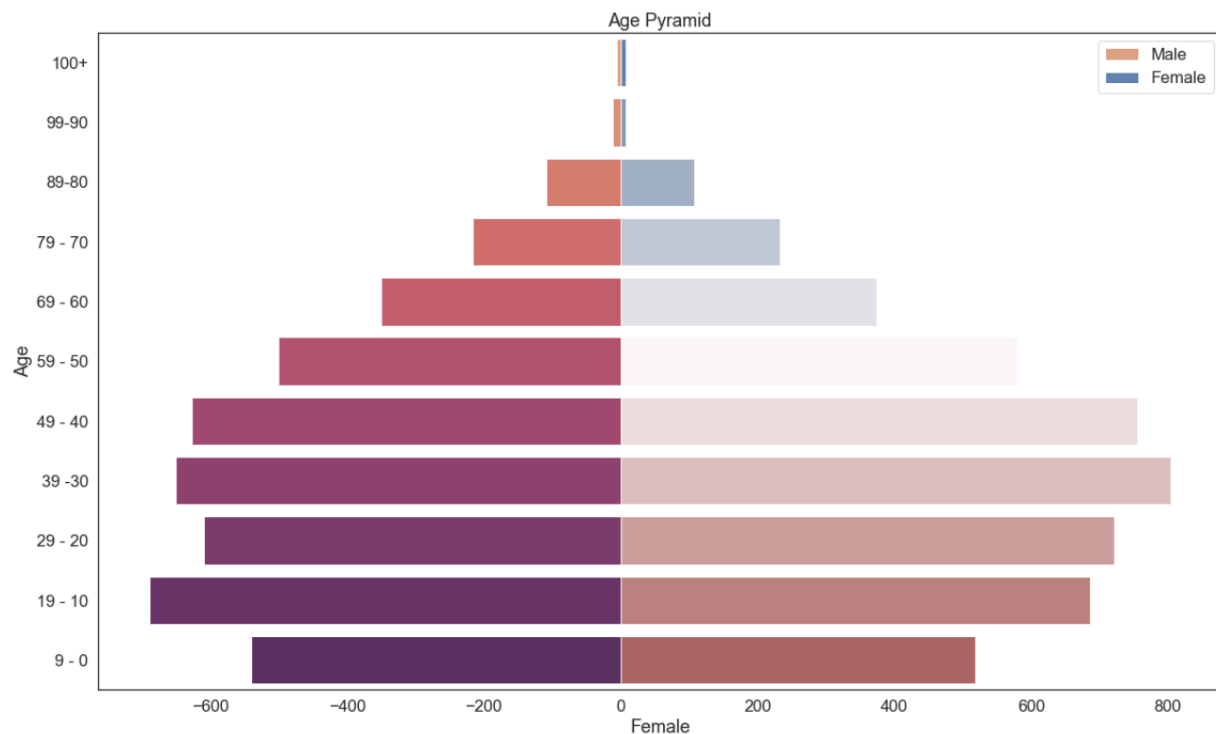
As seen below, the descriptive analysis shows majority of the population are employed, 19.3% of the population are Student, and 6.7% are university student while 5.1% are children which means there is a high number of student with school age. 6.2% of the population are unemployed.

Also, the population of those with 'Single' marital status is higher compared to those who are married and there are more females compared to Males. Furthermore, it is clear that the predominant religion is 'Christianity' and infirmity rate is very low:

Employment

Value	Count	Frequency (%)	
Employed	4878	53.6%	
Student	1755		19.3%
Retired	791		8.7%
University Student	614		6.7%
Unemployed	568		6.2%
Child	466		5.1%
PhD Student	27		0.3%

Religion

Value	Count	Frequency (%)	
Christian	3647	40.1%	
None	3096	34.0%	
Underaged	2138	23.5%	
Muslim	116		1.3%
Sikh	58		0.6%
Jewish	41		0.5%
Bahai	2		< 0.1%
Hindu	1		< 0.1%

Marital Status

Value	Count	Frequency (%)
Single	5469	60.1%
Married	2433	26.7%
Divorced	840	9.2%
Widowed	357	3.9%

Gender

Value	Count	Frequency (%)
Female	4793	52.7%
Male	4306	47.3%

Infirmity

Value	Count	Frequency (%)
None	9048	99.4%
Physical Disability	14	0.2%
Unknown Infection	10	0.1%
Mental Disability	9	0.1%
Deaf	7	0.1%
Disabled	6	0.1%
Blind	5	0.1%

Data Analysis

Religion and Age

The community has a larger percentage of Christians while Jewish, Muslim, Hindu and Sikh also have their worshippers. The population as evident by the median age is not significant enough to build another place of worship. This is further reinforced with the population of people with no religion, or who are underage; both of which are high, thus no new place of worship is required at this time.

Distribution of Age by Religion
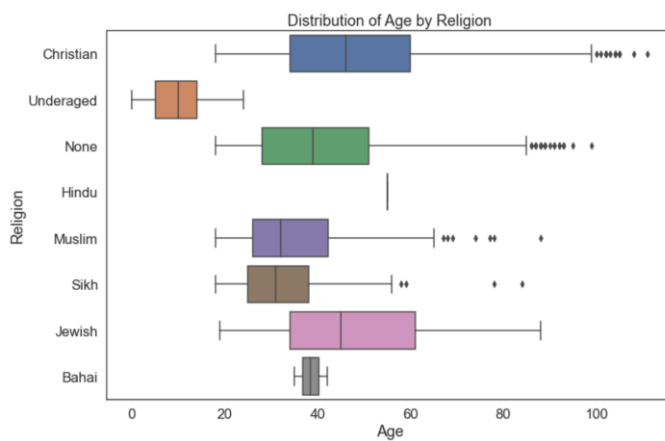
Median Age of Religion

Below shows the median and mean age of each religion:

Religion	mean	median
Bahai	38.500000	38.5
Christian	47.308473	46.0
Hindu	55.000000	55.0
Jewish	49.097561	45.0
Muslim	36.232759	32.0
None	41.296189	39.0
Sikh	34.000000	31.0
Underaged	9.339102	10.0

Divorce and Marriage

```
Groups in Marital Status by Gender:
  Marital Status  Gender  Count
0        Divorced  Female    502
1        Divorced    Male    338
2         Married  Female   1221
3         Married    Male   1212
4          Single  Female   2859
5          Single    Male   2610
6         Widowed  Female    211
7         Widowed    Male    146
```

Following data cleaning, it is immediately apparent that divorce occurs from a young age up until old age, there are also more female divorcees compare to male divorcees as seen above, we can assume some divorcees emigrates from the community and they are most likely males as we have more divorced females compared to male, the difference was used to calculate the emigrants population.

Birth and Death Rate

The birth rates of the population fluctuates, the present estimated crude birth rate is 9.9 births per thousand, the previous 5 years is estimated 11.9 birth per thousand which shows there is a decrease:

```
Present Crude Birthrate per 1000: 9.891196834817013
Previous Crude Birthrate per 1000: 11.869436201780417
Present CBR Change from Previous per 1000: -1.978239366963404
```

The death rate was calculated using age of over 60, there is decrease in the age group of 60 and above, however age 85 has the highest death rate, a further check on the age of 60 and above shows a lot of them are retired, the death rate is 13 per thousand.

```
1  death_rate = abs(((deaths['Count']).sum())/total_population)*1000
2  death_rate
```

```
13.078360259369163
```

Migration

```
100+      -4.0
90-99   -198.0
80-89   -232.0
70-79   -276.0
60-69   -356.0
50-59   -302.0
40-49    -72.0
30-39    124.0
20-29    -45.0
10-19    317.0
0-9        NaN
```

The immigration and emigration in the town includes University students, PhD Students and a certain fraction of divorcees. The above age class shows there is increase in age class 10 -19 and a decrease in 20 – 29 age class which most are students, that means when students graduates, their accommodation will be let out every year. Therefore, students should be considered as part of the community growth. Immigration is calculated from lodgers that are singles and visitors while emigration is calculated from a difference of male and female divorcees, this gives 29.59 immigrants per one thousand while emigrations give 18.02 emigrants per one thousand.

Population Change was also calculated (immigrants_by_1000 + crude birth rate)-(emmigrants_by_1000 + death rate) which makes us conclude the overall population is growing as we have 8.38 even though there is a decrease in growth rate using (crude birth rate - death rate).

Employment and Commuters

The occupation was split into Employed, Unemployed, Retired, Student, University Student, PhD Student and child. The percentage for Employed is 53.6 % which has the highest percentage in the category, using unemployment rate formula, we got 10.42% which is high as the unemployment rate in UK is 4.3% (Trading Economics, 2021). A further check using the age class shows that more people in age class 30-39 and 40-49, are unemployed compared to other age class.

```
30-39     146
40-49     145
50-59     101
20-29      95
60-69      64
70-79       9
80-89       7
10-19       1
0-9         0
90-99       0
100+        0
Name: Age Class, dtype: int64
```

A random sample of 1883 records were picked from the employment bracket to check for assumed commuters which gives us 20.69 % are commuters. Commuters are defined to include University students, PhD students, Lecturer in higher education, e.tc

Occupancy Rate

	House Number	Occupancy Rate
mean	42.043411	2.821395
median	26.000000	2.000000
var	2217.120323	4.028885
std	47.086307	2.007208

Above shows the median occupancy rate of the community which is 2.8, we can round it up to 3 to get the exact number of people, this is used to determine if a particular household is overcrowded. According to [ONC, 2020] the average household size in the UK is 2.4, the mean occupancy rate of this town is is higher and can be used to determine that there is over occupancy in the town. A further check shows 923 houses are overcrowded using a baseline occupancy rate of 3 to check. This check shows that 52% of houses in the town is overcrowded. More so, using the lodgers and visitors to calculate the type of house needed gives us a lower number of people that needs high density housing.

Recommendation

Given the occupancy rate of the town, the town is overcrowded. There is also a potential for families who may need to downsize due to the divorce rates. Therefore, investing on low density housing in the town should be a priority. This will go a long way to reduce the strain on current housing situation and ultimately reduce overcrowding.

Furthermore, there is a dearth of employment for people within the age class 30-39 and 40-49. There is a potential loss of income as a result of unemployment. This age class is within the age that is deemed able to work, thus, it would be prudent to create job centers or vocational centers which will be useful in retraining those who fall into this category. This will ultimately increase employability and income within that group.

Concerning the overall age range, although the birth rate is low at the time, more people are within the middle age group and fewer people in the elderly age group. Although infirmity rate is quite low at the time, based on the current age trend, it is assumed that more people in the middle age group will live long to old age. Therefore, it is not out of place to plan towards developing care homes and hospitals to further promote health as these people are predicted to live to old age.

Finally, there is an increase in commuters within the town, in the long term, building a train station would be recommended to ease the stress off the roads. The population is a young and growing population with a potential to expand rapidly, therefore, having a train station to handle to potential commuter stress should be considered.

References

Why you shouldn't write 'Jedi' as your religion in the Census

Available online:

https://humanists.uk/2021/03/12/why-you-shouldnt-write-jedi-as-your-religion-in-the-census/

[Accessed 06/12/2021]

Parental consent

Available online:

https://www.parliament.uk/about/living-heritage/transformingsociety/private-lives/relationships/overview/lawofmarriage-/ [Accessed 06/12/2021]

Marital status in England and Wales

Available online:

https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationestimatesbymaritalstatusandlivingarrangements/2019#marital-status-in-england-and-wales [Accessed 06/12/2021]

Moving out

Available online:

https://www.nspcc.org.uk/keeping-children-safe/in-the-home/moving-out/ [Accessed 06/12/2021]

Population Change

Available online:

https://www.slideshare.net/kingcobra2012/ib-ess-human-population [Accessed 07/12/2021]

Unemployment rate in UK

Available online:

https://tradingeconomics.com/united-kingdom/unemployment-rate#:~:text=The%20unemployment%20rate%20in%20the%20UK%20increased%20to,the%20coronavirus-hit.%20Figures%20were%20in%20line%20with%20forecasts. [Accessed 08/12/2021]

What is your marital or civil partnership status?

Available online:

 https://census.gov.uk/ni/help/help-with-the-questions/paper-questions-help/what-is-your-marital-or-same-sex-civil-partnerships-status [Accessed 08/12/2021]

Retirement Age

Available online:

(https://www.gov.uk/state-pension-age) [Accessed 08/12/2021]

Population Change

https://slidetodoc.com/population-dynamics-apes-year-end-review-global-population/

[Accessed 08/12/2021]

Families and Household in Uk

Available online:

https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2020

[Accessed 09/12/2021]