Report of Population Census Conducted in a UK Town in 1881

Data Cleaning, Analysis & Recommendations for the Government Officials

December 2020

Prepared and submitted by: OYIN AKINTOYE

Table of Contents

Table of Figures

1.0 Introduction

This is a detailed report of a census conducted in 1881 for a modestly sized sandwiched town between two much larger cities in the United Kingdom. The dataset from the mock census has been analysed to make recommendations for the local government officials in the town to decide on what to be built on an unoccupied plot of land and the most appropriate investment venture for the residents of the town. This report also covers relevant information on the demographics of the population.

2.0 Data Preparation and Cleaning

This section contains the data exploration steps for the dataset, and the cleaning approach used to have a cleaned dataset. The isnull() function in python was used to check for missing values and all other types of missing values presented as blanks, white spaces were all assigned as missing values. The dataset contains no duplicates, has mixture of multiple data types, wrong categorization, value mislabeling, and 10 columns are associated with these data errors.

There are two records with age 17 and recognized as head of household exceptions were made for this assuming they happened to get married early, and I left the records as they will not skew my analysis if removed or adjusted.

```
The mode of the age data is : 19 years
median of age data series: 35 years
mean of age data series: 36 years
The standard deviation of the age data series is: 22
```

Figure 1. A picture of the Descriptive Statistics of Age in the Population

	House Number	Street	First Name	Surname	Age	Relationship to Head of House	Marital Status	Gender
1451	21	Keeper Street	Rebecca	Wright	103	Head	Single	Female
1702	50	Mills Brook	Terry	Owen	103	Head	Married	Male
1703	50	Mills Brook	Kirsty	Owen	106	Wife	Married	Female
2741	51	Haymarket Street	Brandon	Perkins	105	Head	Married	Male
2742	51	Haymarket Street	Nicole	Perkins	106	Wife	Married	Female
2892	14	Exetercrib Avenue	Jasmine	Perry	102	Head	Married	Female
2893	14	Exetercrib Avenue	Gerald	Perry	104	Husband	Married	Male
3058	9	Carter Lane	Mark	Barker	105	Head	Married	Male
3059	9	Carter Lane	Valerie	Barker	108	Wife	Married	Female

Figure 2. A Figure of a Table Showing Outliers in Age of the Population

2.1 NANs

The NANs in my dataset were replaced with the mode of the variable where applicable. For instance, The NaN in my infirmity variable was replaced with the mode of the infirmity status of the population which was then recategorized into non-disabled and disabled if you have a physical or mental impairment that has a substantial and 'long-term' negative effect on your ability to do normal daily activities. (Equality Act 2010). The same approach was applied to Occupation and regrouped into Unemployed, Retired and Employed.

To treat the blank in Last Name, this was replaced using inference from the information of other people in the same household, the blanks were replaced appropriately. I reclassified the ages of the population to Child (<= 17), Adult (18-65), and Senior (65 and above) which will be used for further analysis.

There is a missing gender case for a Head of Household age 58, Single and with a son aged 41. The inference was made to the First Name: Mary to classify her as a Female and the mode of the gender population is Female

Nope, Agnostic, Jedi, Pagan and Housekeeper were not found to be a type of religion in the United Kingdom after several web searches and the recognized religions in United Kigdom were kept (www.ons.gov.uk). All of these were reclassified as None as they could not be associated with any religion currently although there is a likelihood, they take up belief in the future. The NaNs under this column which applies to children and some adults were also replaced with None as the children do not have any religion it is logical to replace the NaNs for adults with None which is also the mode for religion.

2.2 Multiple Data Types

House Number contained a string variable "Two" and was replaced with a numeric "2". Ages are usually expressed as a whole number and if in decimals it is usually rounded off to a whole number. In the dataset, some ages are presented in float (e.g., 79.77226286), to clean this, the entire age column was first converted to a float and then converted to an integer (e.g., 79.77226286 becomes 80). The string values found are converted from text to numeric. (e.g., 'twenty eight' to 28).

I cleaned the single blank in the age column by checking the information from other columns which shows the person's occupation to be a child in a household, Marital Status as NaN which means this person's age is likely to be below age 18 and a child, linking this to information about the census that minors (below age 18) have their Marital Status as NaN. I cleaned this by replacing the missing value by calculating the mode of those below age 18 which is age 17.

2.3 Misspelling and Coding Errors

Due to imputation error, the Relationship to Head of House had a value label misspelling "Neice" and was corrected to Niece. The varying value labels for Marital Status were properly reclassified to (Married, Widowed, Single, and Divorced) and also same was done for Gender.

*Additional variables and dataframes were created by modifying the dataset using several conditions.

3.0 Data Analysis and Detailed Findings

3.1 Population Demographics

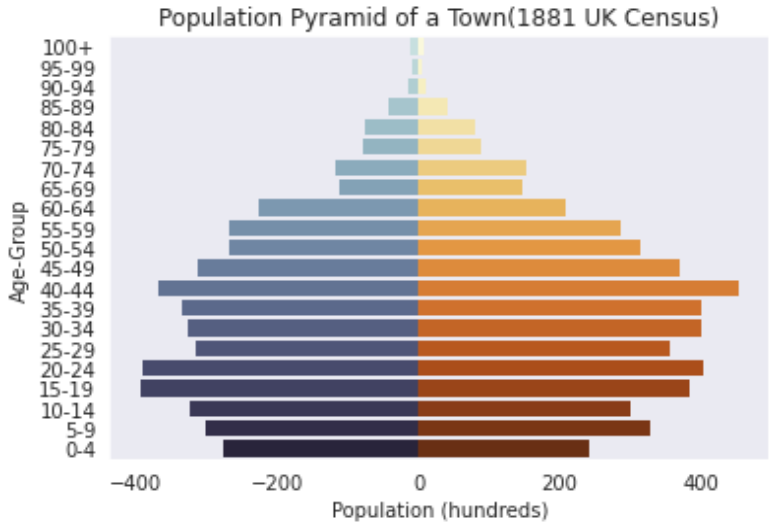

Figure 3. Population Pyramid of a Town in UK in 1881 Census

1. The average wide base indicates a low birth rate as fewer numbers of children are born. Therefore, there is a relatively low birth rate.
2. A narrow apex indicates a high proportion of people living longer although the mortality rate for men is slightly higher than that of women amongst the elderly.
3. The population has many economically active youth (15 - 65 years). From the indents and bulges in the pyramid, there appear to be many activities occurring within this age group such as migration and possibly deaths due to sickness or famine.
4. A high number of young dependents for those below the age of 15. for the age group between 25-29 to 35-39 which are youths and indents shows a high level of migration activities within this age group.

3.2 Birth Rate and Death Rate

To calculate the birth rate by the number of women within the ages of 25-35 as childbearing age for my analysis and those categorized as a child in data are all between age 0-4 because most women in 1800 marry within age 25. The current birth rate of the town is 4.4 per thousand population and the death rate is 5.07 per thousand population.

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

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

Figure 5. Birth Rate and Death Rate Formula

$$\text{Rate of Natural Increase Formula:}\quad \frac{\text{Birth Rate} - \text{Death Rate}}{10}$$

Figure 4. Natural Increase Formula

Furthermore, I calculated a natural increase rate to determine whether the population is increasing or declining without considering migration. A negative rate (-6.70) depicts that more persons in the population are dying than births and indicates a declining population.

The decline in births can be linked to migration to cities, education, and an increase in female employment. A good way the population of the town can grow will be if there was as high immigration and increase in births.

The doubling time calculations show that it will take the town 10.5 years to double its current size if the current growth rate stays the same and all other factors remain constant in the town.

3.3 Housing Occupancy and Density

The total number of households in the town is 3,342 with an average of 4.43 people living in a household. The mean size of occupancy per household is slightly lesser than the mean household size in England and Wales which in the 1880s as a whole was relatively constant at 4.75 (Laslett, 1969).

The proportion of households with occupants higher than the national household size of the United Kingdom is 47% and a little more than half of the households are occupied by persons who are single (53%), followed by those who are married (38%) and 28% are divorced.

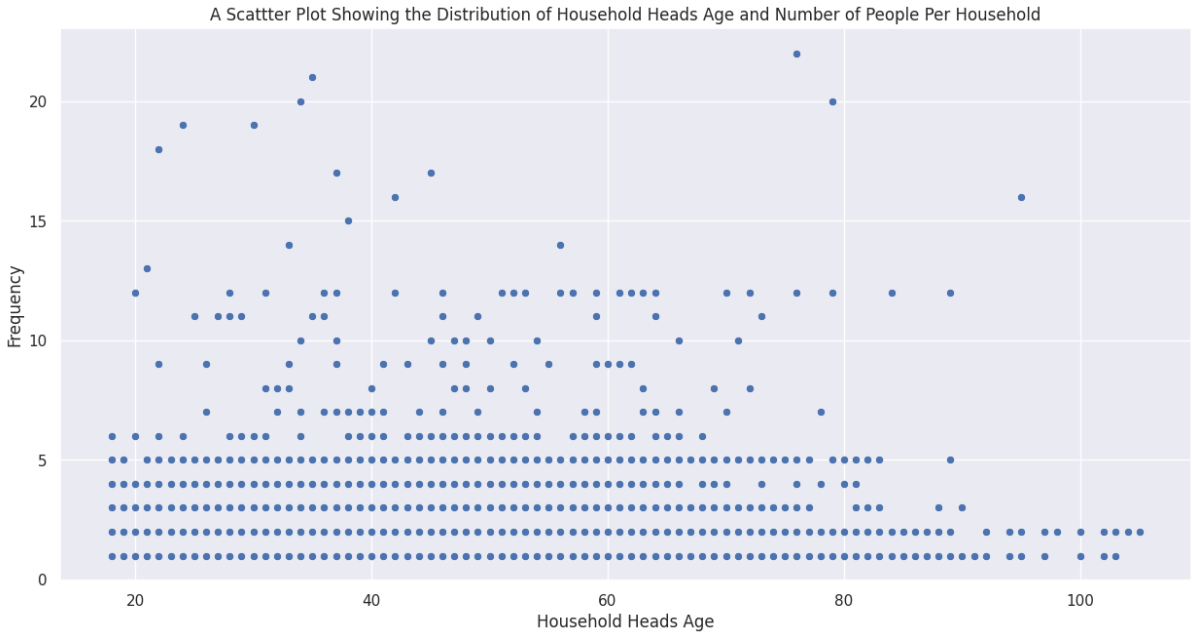

Figure 6. Distribution of Household Heads Age and Number of People Per Household

From the plot below there are extreme occupancy rates that are above 10 with a maximum number of people in a household as 22 and the potential reasons for this could be more people renting their homes to lodgers, those who are into agriculture might have their farm workers living with them and this shows that the government needs to invest in housing as there is likely more lodgers will move into the town, the single people will get married and have more children.

3.4 Marriage and Divorce

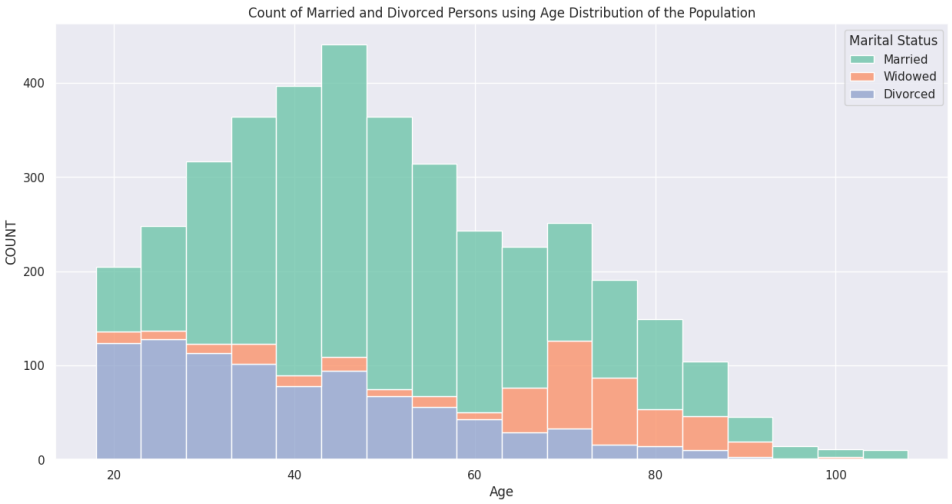

Figure 7. A Chart Showing the Count of Married and Divorced Persons using the Age Distribution of the Population

Marriage ages for men, as indicated by the Singulate Mean Age at Marriage (SMAM), rose gradually throughout the period, from 26.1 in 1851 to reach 26.6 by 1881, 27.2 by 1901, and a high of 27.6 by 1911. In comparison, the SMAM for women fell slightly between 1851 and 1861, from 25.5 to 25.2, then remained level before increasing to 25.8 in 1891 (Schürer et al., 2018).

The measurement of refined divorced rate puts the number of divorces among married persons is at 35% per 1,000 women in the population and from the chart above it appears the divorce rate is most common amongst young people and will be great if counselling services are provided to keep people married by the government.

3.5 Employment and Training

My analysis categorises the economically active population as 16 - 64 years (Office for National Statistics, 2018). Comparing the unemployment rate (4.91%) in the United Kingdom in 1881 to the town's unemployment rate of 9.6%, the unemployment rate is very high (Bank of England, 1760). A deep dive into the analysis of those unemployed and who should be working shows that the majority of these persons are in their early twenties which depicts they are most likely young school leavers who have a good chance of later getting a job although the available job opportunities might be outside town. As such there will be no need to invest in employment, training or retraining.

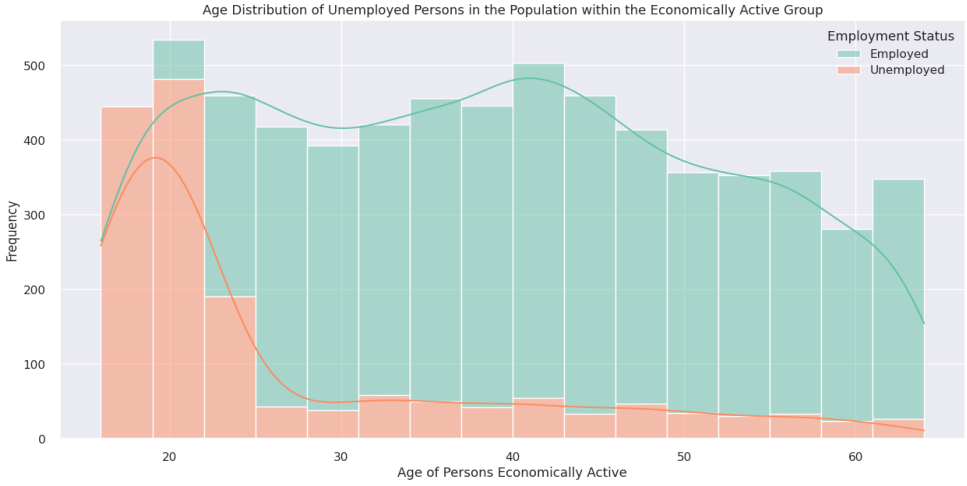

Figure 8. Chart Showing Age Distribution of Unemployed Persons within the Economic Active Population

Demographic labour pressure index (DLPI) is calculated as follows:

$$DLPI = \frac{\sum_{5}^{14} R}{\sum_{55}^{64} R},$$

Figure 9. Demographic Labour Pressure Index Formula

The current Demographic labor pressure index is 1.0 which depicts that the ratio of persons who will enter the labor force in the next 10 years to the number of persons that will exit will be the same.

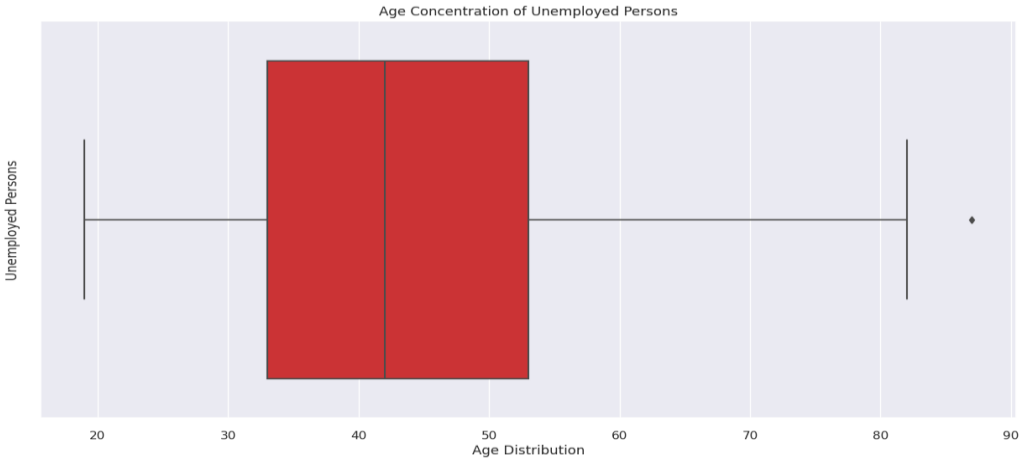

Figure 10. Chart Showing Age Concentration of Unemployed Persons in the Town Population

3.6 Commuters and Migration

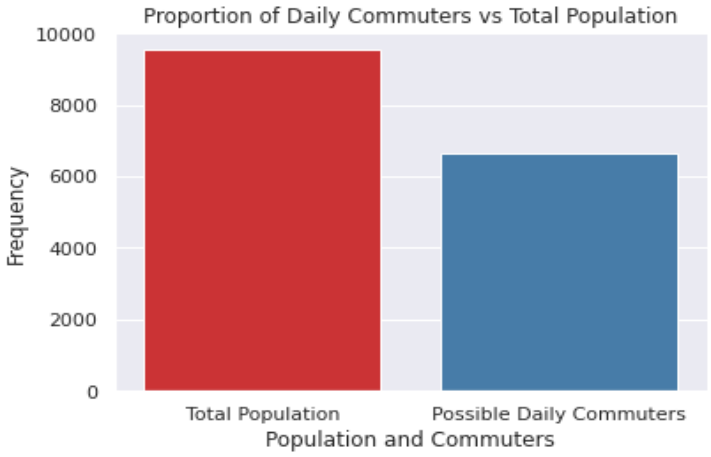

Figure 11. Chart Showing the Total Population and proportion of Commuters in the Population

Those in the population with professional employment and students are assumed to be commuters while occupations that are likely to be for those self-employed such as restaurant, retail, shop owners, unemployed and those retired were taken to be non-commuters. A large proportion of the population are discovered to be commuters (70%) as such a good and improved transport system such as a train station will be of great benefit to the populace and can attract more immigrants.

3.7 Religion

Across all the age groups those with *None* religion are the majority of the population although this declines as people get older. The chart below clearly reveals that Christians are a growing religion and are the leading denomination in the population amongst those who are religious but most of their followers are highly concentrated within the adult population (40-60 years).

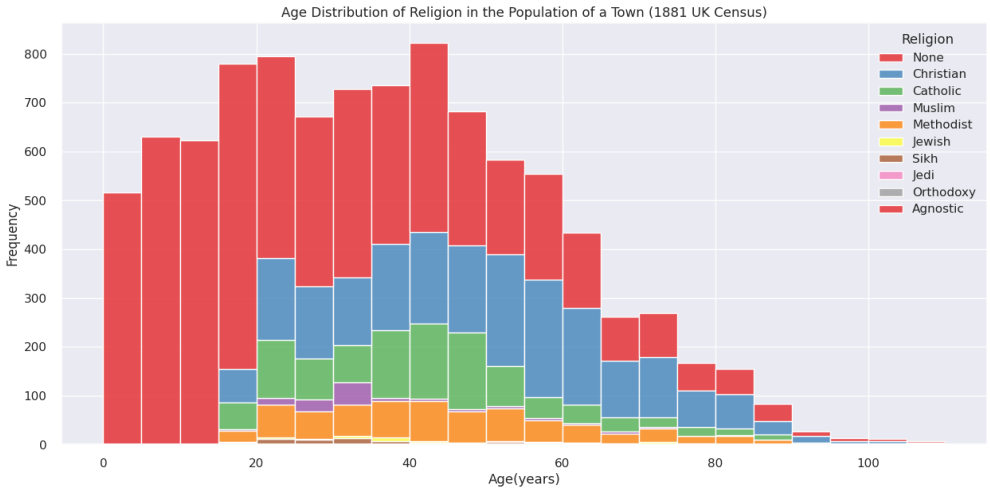

Figure 12. Chart Showing the Age Distribution of Religion Types in the Population

All other religious types have a very small growth and do not have a sharp increase and dominant members within the younger population enough to recommend building a religious building for any of them at the current time and perhaps they might grow in the future.

3.8 Infirmity

I split the infirmity group into disabled and non-disabled (GOV.UK, 2011). The population is a very healthy population as only less than 1% of the population is disabled and it is not important for the government to push investment but ensure they keep maintaining a healthy population.

4.0 Recommendation

1. There is evidence of a high unemployment rate in the population amongst the economically active population as such it is deemed necessary to train and retrain people for new skills and more job opportunities provided.
2. For a small town, the occupancy mean is very close to the national household size mean and the town has some excessive occupancy running into twenty-two coupled with a huge portion of its population being youthful and single who will transition into being married and build a family. It is a necessity to build a low-density housing as most of the population are affluent by being gainful employed.
3. The proportion of elderly dependents in the population is low, hence, it is not necessary to invest in elderly care homes, improving healthcare costs and retirement payments. However, this will be necessary in the future as majority of the population who are youthful will start ageing.
4. The doubling time of the population is 10.5 years if the condition remains the same thus it is cogent for the town to plan for its double capacity every 10.5 years for waste management, water supply, road maintenance and other services.
5. Investment in schooling and another religious building is not necessary at this point because the town has a low birth rate, and the thriving religion (Christianity) is mostly the adult population who will soon be retired.

5.0 References

Bank of England (1760). *Unemployment Rate in the United Kingdom*. [online] FRED, Federal Reserve Bank of St. Louis. Available at: https://fred.stlouisfed.org/series/UNRTUKA.

GOV.UK (2011). *Definition of Disability under the Equality Act 2010*. [online] GOV.UK. Available at: https://www.gov.uk/definition-of-disability-under-equality-act-2010.

Laslett, Peter. "Size and Structure of the Household in England Over Three Centuries." *Population Studies*, vol. 23, no. 2, 1969, pp. 199–223. *JSTOR*, https://doi.org/10.2307/2172902. Accessed 8 Dec. 2022.

Schürer, K., Garrett, E., Jaadla, H., & Reid, A. (2018). Household and family structure in England and Wales (1851–1911): Continuities and change. *Continuity and Change, 33*(3), 365-411. doi:10.1017/S0268416018000243

Office for National Statistics (2018). *Long-term trends in UK employment: 1861 to 2018 - Office for National Statistics*. [online] Ons.gov.uk. Available at: https://www.ons.gov.uk/economy/nationalaccounts/uksectoraccounts/compendium/economicreview/april2019/longtermtrendsinukemployment1861to2018.

Schürer, K., Garrett, E.M., Jaadla, H. and Reid, A. (2018). Household and family structure in England and Wales (1851–1911): continuities and change. *Continuity and Change*, 33(3), pp.365–411. doi:10.1017/s0268416018000243.

www.ons.gov.uk. (n.d.). List of recognised religions in England and Wales - Office for National Statistics. [online] Available at: https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/listofrecognisedreligionsinenglandandwales.