

INTRODUCTION

Background and Purpose of the census

This is a fabricated census that was generated to simulate a census carried out in the United Kingdom. The purpose of such a census is to compare the different people across the nation and to provide the government with accurate statistics of the population to enable better planning, to develop policies, and to allocate certain funding.

The scope and coverage of this census is an imaginary modest town. And its main importance to be able to at the end come up with a decision on what to do with an occupied plot of land and what to invest in

METHODOLOGY

Sampling Design

The sampling design for the successful execution of this census was to distribute a team to go to every street in the town and to go to each house to collect information.

Data Collection Procedures

The data collection procedure for this census was, for every house in which successful entry was gained, a set of the same question were asked of each household. In which after the data was compiled together into a csv. file

Data Processing And Analysis

After successfully gathering and compiling the data, the data was cleaned. Activities carried out while cleaning the data include;

- Dropping of a column that went by 'Unnamed:0'
- Creating a new column called 'Occupation Category', which is a category generated from the 'Occupation' column that categorizes the Occupations into 5 major categories which are; 'Employed', 'Unemployed', 'Student', 'Child', 'Retired'.
- Dropping Exactly 6 rows due to the fact that they contained dirty data.
- Reducing the Number of unique values in the 'Marital Status' Column by renaming
- Renaming of 'Marital Status' and 'Relationship to Head of House' columns to 'Marital_Status' and Relationship_to_HOH' respectively
- Filled in some of the null values in 'Marital_Status' and 'Religion' column using KNNImputer, while dropping some in order not to encourage bias.
- Creating an 'Age group' column that categorizes the 'Age' column into 5 categories which are; Youngins - Individuals less than 11
Gen Z - Individuals in between the ages 10 and 27
Millennials - Individuals in between ages 26 and 43
Gen X - Individuals in between ages 42 and 59
Boomers ii - Individuals
- Reduced the Gender categories to two main categories 'Male' and 'Female'
- Created an 'Address' column created by joining the house number and street columns together.

After the cleaning of the data was carried out, the data was then visualized.

Data Visualization

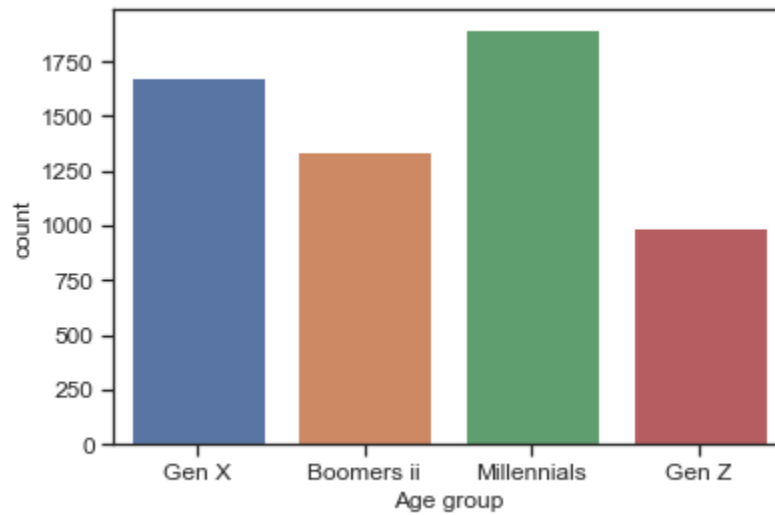


Fig 1 (Barplot showing the age group distribution amongst the population)

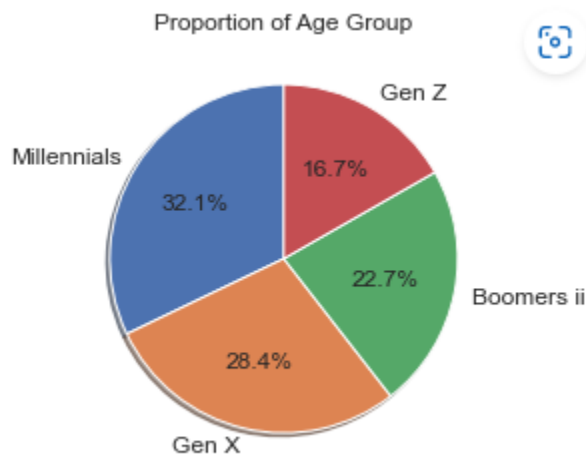


Fig 2 (Pie chart depicting the age group distribution in percentage)

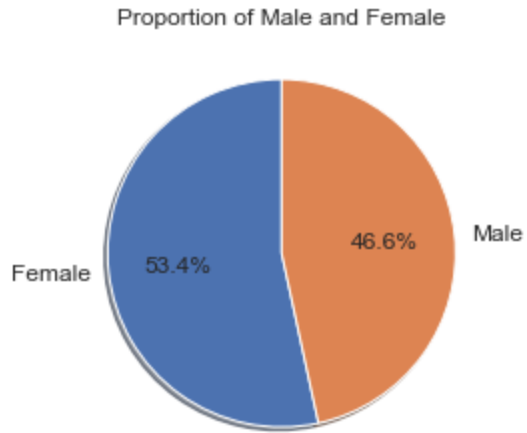


Fig 3 (Pie chart showing the population distribution of the genders)

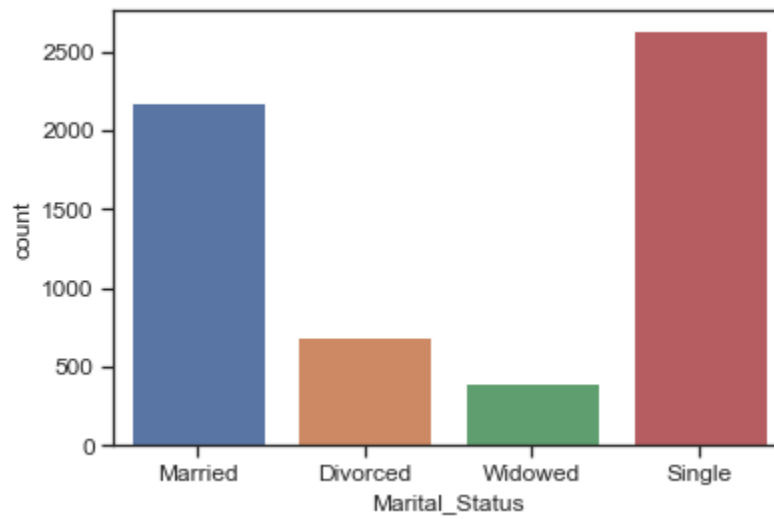


Fig 4 (Bar plot showing the the marital status distribution amongst the population)

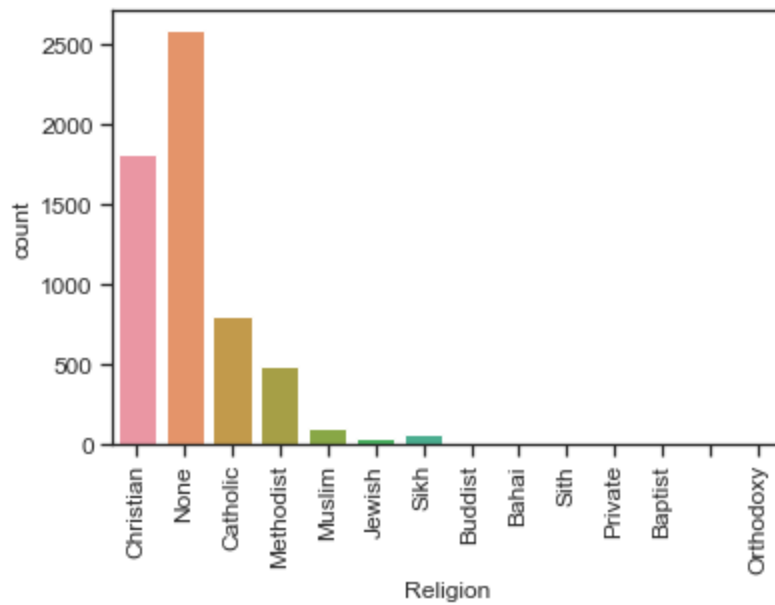


Fig 5 (Bar plot showing the religion distribution in the population)

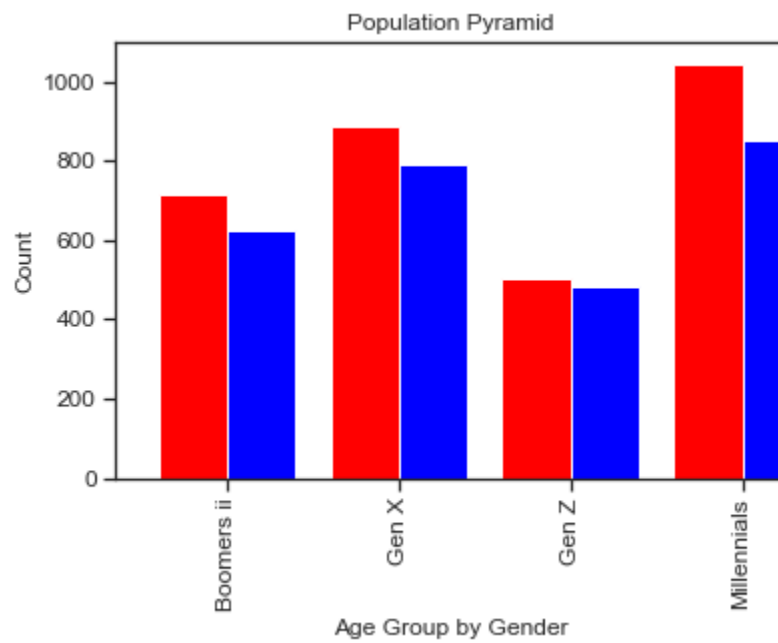


Fig 6 (Bar plot showing the distribution of the age groups in relationship with gender)

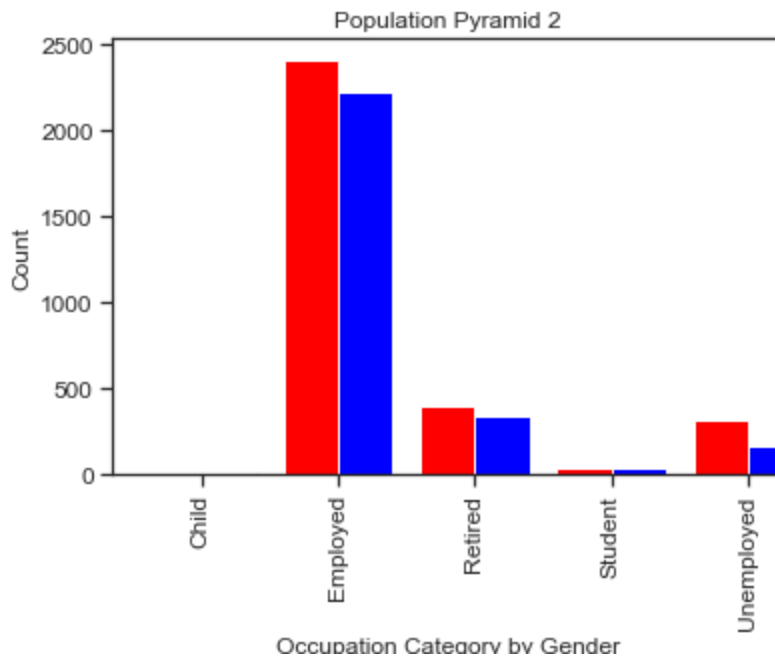


Fig 7 (Barchart showing the Occupation distribution in respect to gender)

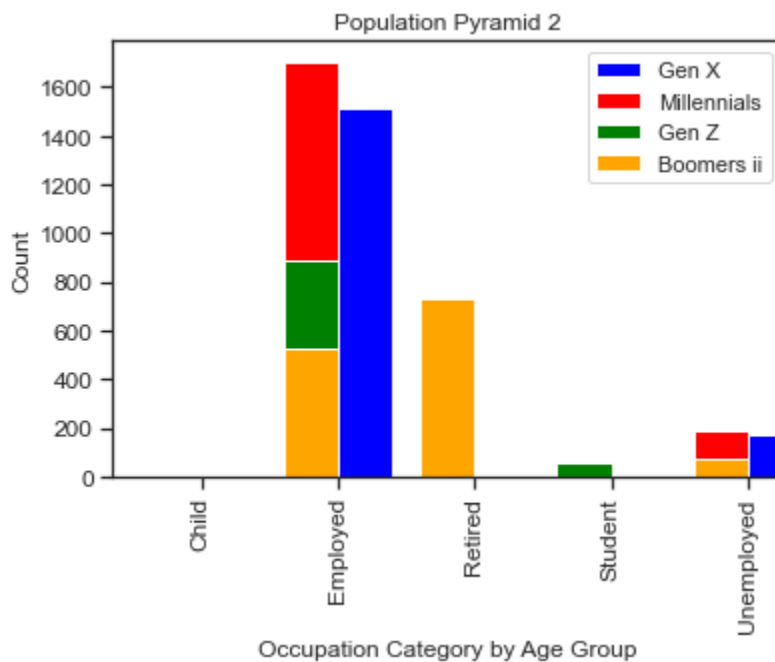


Fig 8 (Histogram showing the occupation distribution in relation to the age groups)

Interpretation

From the findings, the interpretation of the above presented from the data and its visualizations are as follows.

- The general population consists of both male and females with females of the majority and this can be observed in Figures (3, 6 and 7)
- The general population consists of significantly of four age groups which are Gen Z (10-27), Gen X(42-59), Millennials(28-41) and Boomers ii(60 and above) and the Millennials have the highest population count while Gen Z have the lowest counts amongst the distribution
- The marital status of the general population comes in 4 categories which are single, married, divorced and widowed. From the data visualization from fig 4, it is clearly observed that the singles take up the majority of the population followed by the married individuals while the widowed are the minority.
- As Stated earlier, the female gender takes the majority of the population, and because of this the female gender also takes up the higher percentage when it comes to occupation and the age group categories according to Fig(7 & 6) respectively.
- In the location in which the census was carried, the population is made up of different religion sectors which are(Christian, Catholic, Methodist, Muslim, jewish, sikh , buddhist, bahai, sith, private, baptist, orthodoxy and finally None, which are simply people that don't practice any of the aforementioned religions). And according to the data presented and visualized in Fig 5, The people categorized as None have the highest percentage amongst the populace, followed by the Christians and then with a very significant difference between them the catholics come next and after which the methodist. The remainder of the religions occupy a very small ratio of the population.
- In the population, the occupations are categorized into 5 major groups which are; Child, Student, Employed, Unemployed and Retired. According to Fig 8, the majority of the population is mostly employed with the millennials occupying a majority of the employed individuals followed by the Gen X, Gen Z then Boomers ii respectively. After which the retired individuals take the next position with all of them being in the boomers ii category. Then the students come next and they take a very small ratio of the population. Then lastly the unemployed individuals which occupy a very insignificant ratio and are occupied by all age groups except Gen Z.
- And finally, the amount of occupants per housing was generated and the average was deducted by calculating the mean and median which are 3.0 and 2.0

respectively. This simply means that an average of three individuals occupied most houses while houses with 3 occupants were the most common.

Conclusions

After a census was carried out, it was shown with the use of data generated that the population is composed of both male and females with the females having the higher ratio by less than 10 percent. All age groups are present but the millennials are the most while the Gen Z are the least. Populated by various religions the Christians have the most number after the individuals categorized as None. And finally the population consists of mostly the working class from all age groups.

Task 1

1) What should be built on an occupied plot of land that the local government wished to develop.

Options are but not limited to;

- **High density housing - If population is significantly expanding**
- **Low density housing - If the population affluent**
- **Train station - If they are a lot of commuters**
- **Religious building - There exists only one worship center and it is for catholics.**
- **Emergency medical building - If there are many injuries or future pregnancies.**

Answer

Due to the fact that we were not presented with data regarding birth rate or mortality rate, we can not ascertain as to the expansion of the population of the town, the income of the citizens was also not generated so it can be said if they are affluent, average or below average. And also the majority of the housing contain an average of 3 occupants which can be said to be a moderate use of occupancy. But the city does have a significant amount of employed individuals which make up the majority of the population and building a train station will not only assist in making commuting faster but it will also take pressure off the roads. Therefore, it is recommended the local government develop a train station.

Task 2

1) Which of the following options should be invested in.

Options include:

- **Employment and training**
- **Old age care**
- **Increase spending for schooling**
- **General Infrastructure.**

Answer

It is in my opinion that Investments should be directed towards old age care simply because the employed population takes up most of the percentage in comparison to the whole population and that niche is mostly occupied by the Gen X, Millennials, Gen Z and Boomers ii respectively and also after the employed population, the retired individuals have the next most count. And in the nearer future, the boomers ii would be retired and the other aforementioned age categories would be also retiring or at least getting close to the retirement age. So in order to prepare for this, investments in old care should be priority