CENSUS 13 PROJECT REPORT

A moderately sized town's census was analyzed in this report to provide recommendations for spending on upcoming services using cases to suggest developments on the unused piece of land.

The census data was first cleaned to address data inaccuracies and missing entries before these suggestions could be made about the investment projects.

In this report, emphasis will be made on important analyses that were conducted particularly to support the suggestions made. This comprises a summary of the demographics of the town's population, age pyramid, followed by a thorough study of the expected increase of the town's population, job patterns, commuters, and occupancy rates.

DATA CLEANING

Before these recommendations on the investment projects could be generated, the census data had to be cleansed to remove errors and missing items as shown in the Jupiter notebook.

When searching for null values, a large number of missing values in the columns for marital status and religion were found.

The null values for children from age 0 to 17 who have not attained the marriage age were replaced with **NA** and the null values for the **Marital Status** column were changed from null to **Single** for the rest of the population. (According to The Marriage and Civil Partnership (Minimum Age) Act 2022 in the UK).

The null values for the Religion column were replaced with None including children under the age of 18. This is because parents have the legal right to choose a religion for their children. Some unique religions were discovered in the census data e.g. Private and Nope which I converted to None. Furthermore, I decided not to change some Religions with low counts in other not to infringe on their human rights of religion. (Freedom of Religion or Belief (FoRB) by UK government)

The missing value under Gender was replaced by Male considering the relationship to the head of house. Some empty values were also discovered in the surname, relationship to the head of house, marital status, gender, occupation and infirmity columns of the census data. The empty values in the surname column was replaced by **unknown** as there is no relationship with other loggers in the houses. There is a unique member of the town with missing marital status but has husband as the relationship to the head of house, therefore this empty space was replaced by Married as this is logical. The missing gender information was replaced by Male due to the relationship with the head of house.

Missing values in the occupation column was replaced by Retired due to the age and it is believed that most women will retire before the age of 71. Another empty space in the occupation was also replaced by Unemployed because the age (46) is within the working class. Furthermore, the empty spaces in the infirmity column were replaced by none respectively.

DATA VISUALIZATION

Age Pyramid
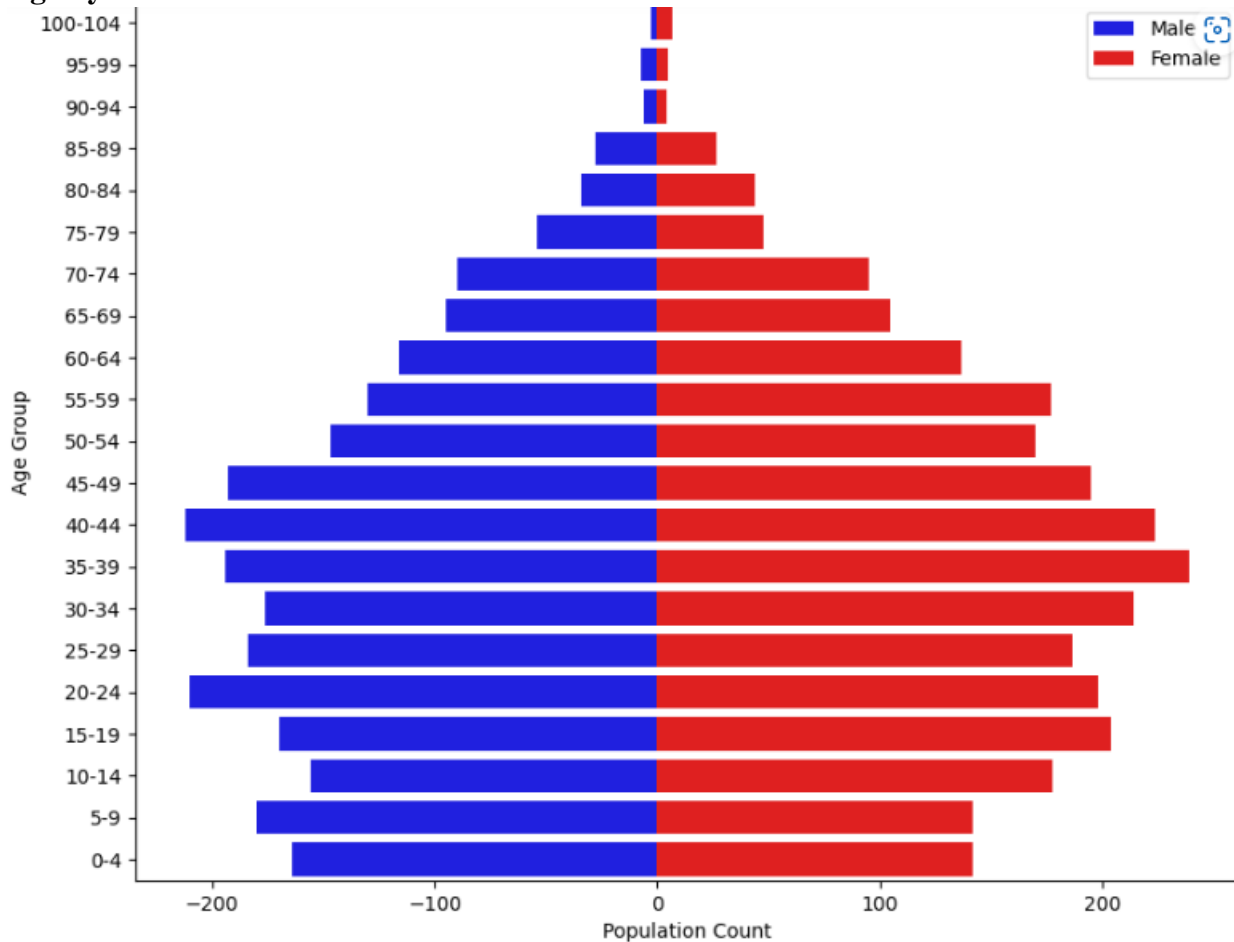
From the Age pyramid above, we can see that the number female population is slightly higher than the male population with increasing rate of growing population. Furthermore, there is a high population of residents of the town from age bracket 20 - 49 who also form part of the university students and working-class population. Furthermore, it is clear from the above age pyramid that the also a growing population of the ages below 20 years.

UNEMPOYMENT TRENDS

The commuters in the town can be identified by

- University students as well as the PHD students.
- Unique occupations that will require travelling in and out of the town e.g Academic Librarian

Unemployed Count by Age
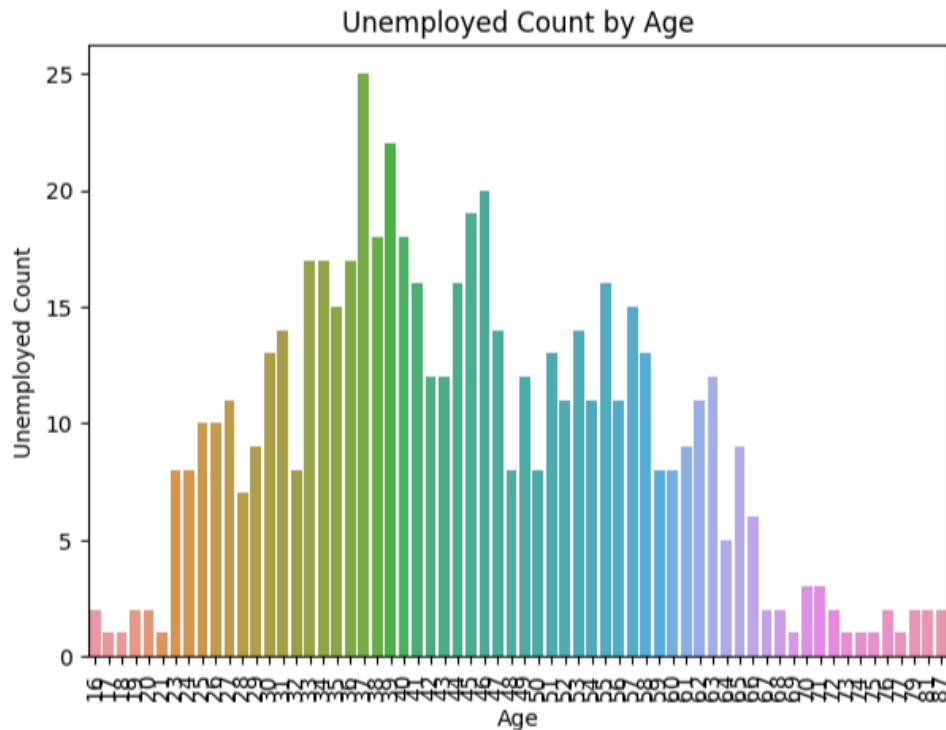
The above diagram shows unemployment within the working age of the population (16-64years according to working age population UK)


```
In [95]: data[['Age','Occupation']].groupby('Occupation').describe()
```

Out[95]:

Occupation	count	mean	std	min	25%	50%	75%	max
Academic librarian	8.0	46.500000	15.371588	23.0	35.50	48.0	57.00	67.0
Accommodation manager	6.0	37.833333	9.867455	26.0	30.50	37.5	44.50	51.0
Accountant, chartered	9.0	49.111111	12.323194	29.0	46.00	49.0	58.00	67.0
Accountant, chartered certified	6.0	44.333333	13.002564	32.0	34.25	42.0	49.75	66.0
Accountant, chartered management	8.0	45.000000	8.585702	32.0	41.50	46.0	49.50	58.0
...	...	...	...	...	...	...	...	...
Water quality scientist	6.0	50.000000	10.119289	35.0	43.25	53.0	56.75	61.0
Web designer	6.0	45.166667	13.392784	31.0	35.25	42.0	54.00	65.0
Wellsite geologist	7.0	40.000000	9.327379	26.0	34.50	40.0	46.50	52.0
Writer	7.0	48.142857	16.035675	29.0	33.00	53.0	62.00	65.0
Youth worker	11.0	44.636364	8.164224	32.0	39.50	46.0	50.50	54.0

1126 rows × 8 columns

```
In [68]: data['Occupation'].value_counts()

Out[68]: Student                                    1577
         University Student                          592
         Unemployed                                  590
         Child                                       494
         PhD Student                                  20
                                                     ...
         Retired Designer, interior/spatial            1
         Retired Gaffer                                1
         Retired Sports development officer            1
         Retired Public relations officer              1
         Retired Animal nutritionist                   1
         Name: Occupation, Length: 1126, dtype: int64
```

The above table shows the variations in Occupation by Age.

RELIGION AFFLIATIONS

Religion Distribution by Age
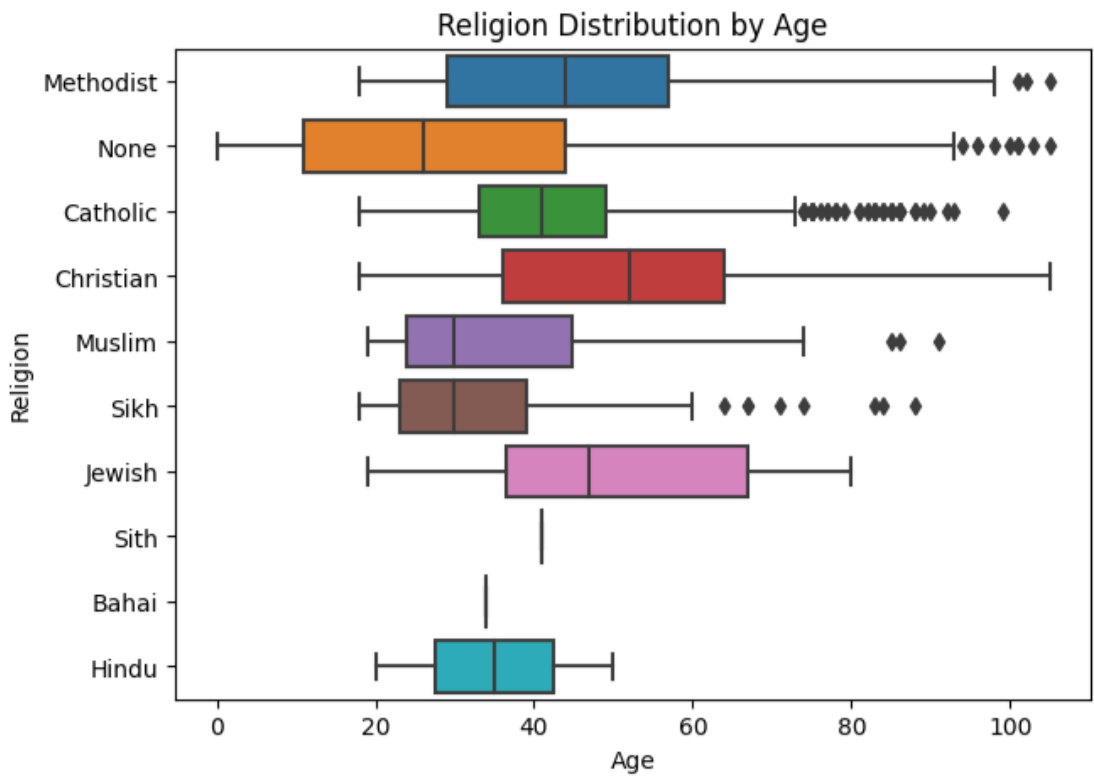
The above plot shows the Religion Distributions by Age

Religion	count	mean	std	min	25%	50%	75%	Age max
Bahai	2.0	34.000000	0.000000	34.0	34.0	34.0	34.00	34.0
Catholic	955.0	42.556021	15.273322	18.0	33.0	41.0	49.00	99.0
Christian	2100.0	50.781905	18.587420	18.0	36.0	52.0	64.00	105.0
Hindu	2.0	35.000000	21.213203	20.0	27.5	35.0	42.50	50.0
Jewish	40.0	50.200000	17.221334	19.0	36.5	47.0	67.00	80.0
Methodist	578.0	44.519031	17.880127	18.0	29.0	44.0	57.00	105.0
Muslim	130.0	36.600000	16.699626	19.0	24.0	30.0	44.75	91.0
None	5005.0	29.395005	21.543016	0.0	11.0	26.0	44.00	105.0
Sikh	65.0	35.815385	18.109815	18.0	23.0	30.0	39.00	88.0
Sith	1.0	41.000000	NaN	41.0	41.0	41.0	41.00	41.0

The above table shows the Religion variations by Age in the town from the census data. It is

obvious that the is a large percentage of Christian in that town. I have replaced 'Private' and 'Nope' responses by None and also decided not to replace some religions with low counts (Sith) based on "Freedom of Religion or Belief (FoRB) which give freedom of religion in the UK.

MARRIAGE AND DIVORCE RATE
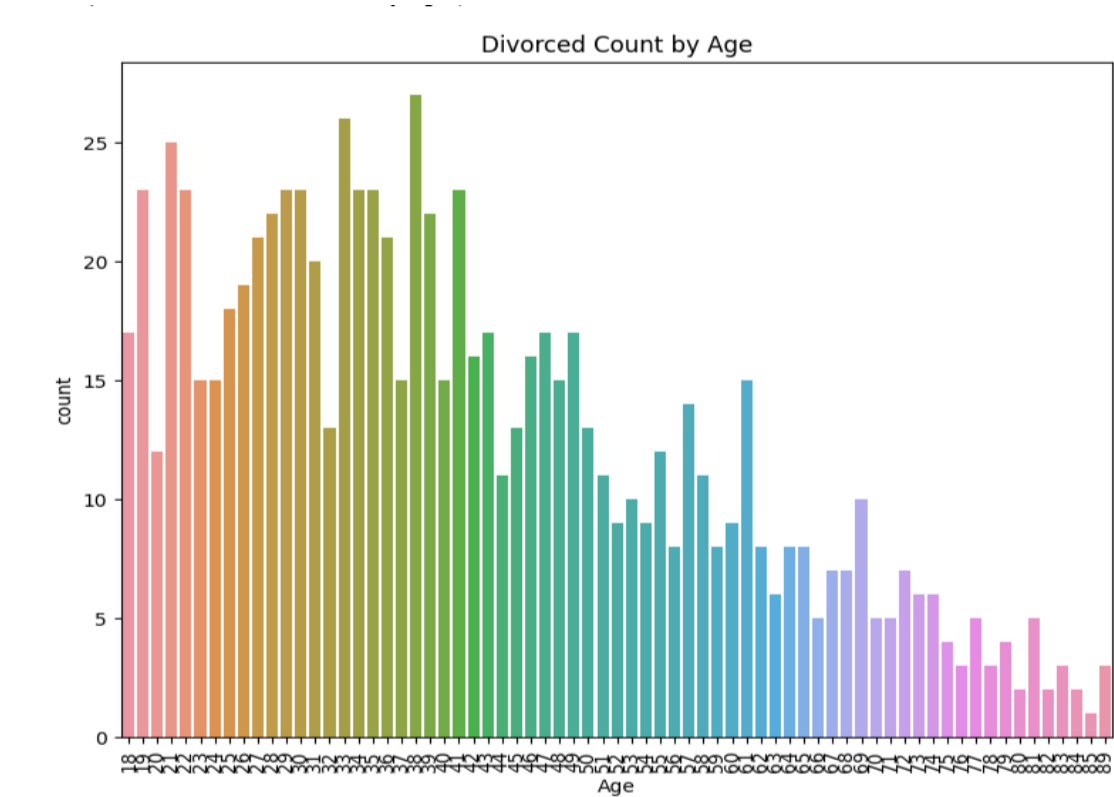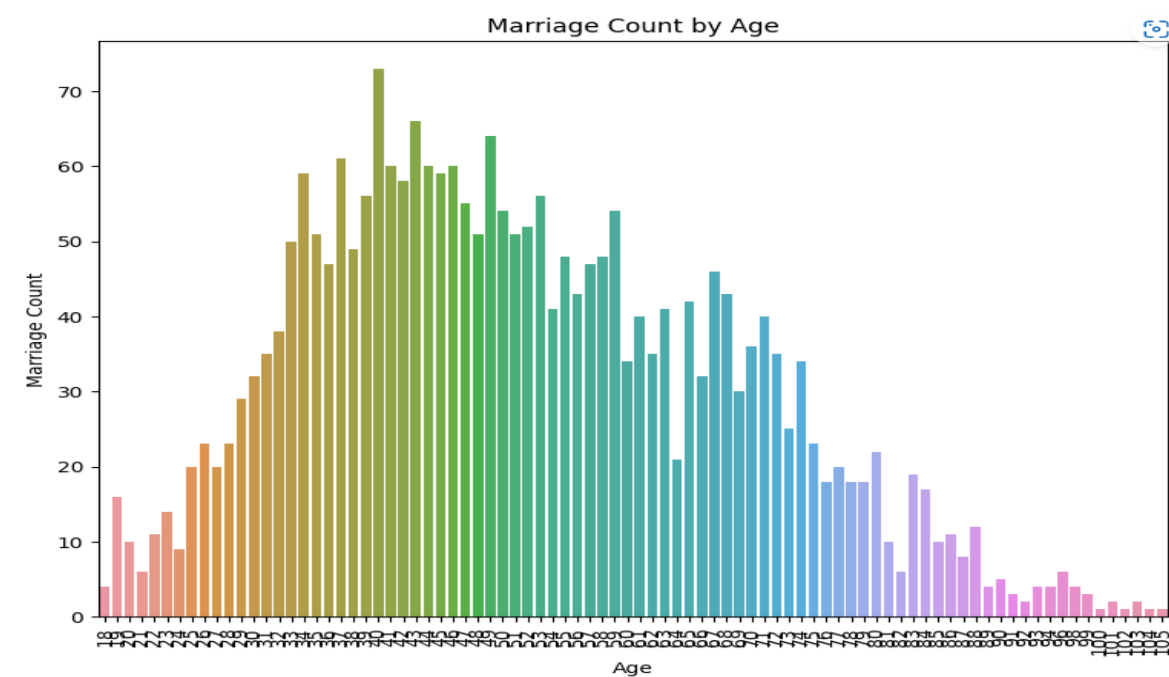
```
In [75]: data[['Age','Marital Status']].groupby('Marital Status').describe()
```

Out[75]:

					Age			
	count	mean	std	min	25%	50%	75%	max
Marital Status								
Divorced	860.0	41.943023	16.575794	18.0	29.0	39.0	53.0	89.0
Married	2552.0	51.704154	16.724315	18.0	39.0	50.0	64.0	105.0
NA	1970.0	8.879695	5.106310	0.0	4.0	9.0	13.0	17.0
Single	3093.0	37.720983	14.339199	18.0	25.0	36.0	48.0	100.0
Widowed	403.0	67.558313	18.395691	18.0	66.0	72.0	78.0	105.0

```
In [76]: # Percentage of Married Population

Percentage_Married_Rate = round(len(Married_Population)* 100/ len(data),2)
print(f"{Percentage_Married_Rate}%")
```

28.75%

```
In [79]: # Percentage of Divorced Population

Percentage_Divorced_Rate = round(len(Divorced_Population)* 100/ len(data),2)
print(f"{Percentage_Divorced_Rate}%")
```

9.69%

The above diagrams show the married and divorced counts across the ages the census data. Population below the age of 18 were replaced by Single (The Marriage and Civil Partnership (Minimum Age) Act 2022). Standard deviation also provides information on the variation in the age distribution of the population across the different categories of Marital Status. The percentage of the married population is 28.75% while divorced population is 9.69%.

OCCUPANCY LEVEL

```
In [80]: Occupancy = data.groupby(['Street', 'House Number']).size().reset_index(name='Number of Occupants')
Occupancy['Number of Occupants'].value_counts()
```

```
Out[80]: 2     966
         1     964
         3     538
         4     511
         5     293
         6      67
         7      12
         12      9
         8       6
         10      5
         11      4
         15      3
         9       3
         13      1
         21      1
         17      1
Name: Number of Occupants, dtype: int64
```

The above table shows the Occupancy level of the population in the Town.

INFIRMITY

```
Out[99]:
```

	count	mean	std	min	25%	50%	75%	Age max
Infirmity								
	8.0	40.625000	16.097360	15.0	30.75	39.0	55.50	61.0
Blind	8.0	34.125000	19.650064	3.0	25.50	32.0	47.50	62.0
Deaf	9.0	38.666667	34.803017	2.0	13.00	20.0	66.00	94.0
Disabled	11.0	31.636364	20.857961	7.0	20.00	27.0	41.00	79.0
Mental Disability	10.0	44.400000	23.557023	3.0	32.50	39.5	56.25	88.0
None	8805.0	37.073367	21.966670	0.0	19.00	36.0	53.00	105.0
Physical Disability	18.0	46.055556	17.027563	5.0	37.75	46.0	53.00	74.0
Unknown Infection	9.0	45.777778	12.132372	26.0	40.00	43.0	57.00	65.0

The above table also show the Standard deviation outlining the variations of infirmities of the population in the town. It can be seen that there is low infirmity rate in the town.

RECOMMENDATIONS

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

The issue of transport infrastructure needs to be addressed due to the large number of commuters who travel daily as well as the growing population of teenagers who are approaching university level. Building a **Train Station** that can safely and effectively transfer lots of commuters to and from their destinations is an ideal recommendation on the unused land in the town.

There would be several advantages for commuters and the nearby communities if a train station were built. For instance, it would offer a dependable and safe mode of transportation, lowering the likelihood of traffic jams and accidents. A train station would also be a more environmentally beneficial choice, encouraging sustainability and lowering carbon emissions.

Furthermore, a train station would also facilitate the transportation of products and services and create new job possibilities, which would assist in strengthening the local economy. As a result, the town would attract new companies, promoting development and economic progress.

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

Investing in **Employment and Training** can be beneficial to the community because of the level of unemployment. These programs can give people the knowledge and training they need to land better-paying jobs, lowering the general unemployment rate and boosting the economy.

Employment initiatives like job training, apprenticeships, and internships can be created to equip people with the unique expertise needed for a given position or industry, which can improve their employability and income potential.

Overall, investing in **Employment and Training** can benefit the whole population and increase the economic developments of the town.

REFERENCES

1. Freedom of religion or belief: understanding this human right Website: Gov.uk https://www.gov.uk/guidance/freedom-of-religion-or-belief-understanding-this-human-right
2. Government of the United Kingdom. (2021, June 30). Legal age of marriage in England and Wales rises to 18. Retrieved from https://www.gov.uk/government/news/legal-age-of-marriage-in-england-and-wales-rises-to-18
3. Working Age population https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/working-age-population/latest