Investigating the demographic and socioeconomic traits of people in UK 2021

1 Introduction

The data set used for this analysis is a modified snapshot of information gathered in 2021 during a UK household census. It includes a range of factors that provide information about the socioeconomic and demographic traits of people and families. Household identification, individual identification, age, marital status, annual income, gender, accomodation type and highest level of education attained are key variables used in this analysis. To do the analysis various R packages has been used like Tidyverse (Wickham et al. 2019) and Ggplot2 (Wickham 2016)

2 Data

This data set extracted from the UK census data 2021. Census in UK happens after every 10 years and gives the snapshot of all the people and households in UK. Conducted at a national scale, the census served the purpose of informing public policy, resource allocation, and social research by providing a detailed snapshot of the population's characteristics National Statistics (2023).

▼ Code

```
data <- data_1 |>
    rename(Marital_Status= Mar_Stat, Income = INC, Gender = Female, Accomodation_Type = H8, Et

data$Accomodation_Type <- factor(data$Accomodation_Type, levels = c(0, 1), labels = c("Self

  data <- data |>
  mutate(Gender = ifelse(Gender==1,"Female","Male"))
```

3 Low Income Households and Individuals Insights

The data contains 10565 households. Out of 10,565 , 6730 (63.7%) households have annual income less than 17,000. It is quite alarming figure. To further investigate this, identification of low income individuals by age bracket, ethnicity, gender has been done. 17,700 figure considered to be in low income category has been decided from this source GOV.UK (2023)

3.1 Low Income Individuals by Age Group

▼ Code

```
data$Age_Bracket <- cut(data$Age, breaks = c(0,18, 30, 40, 50, 60, Inf), labels = c("0-18","19
low_income_data <- data |>
    filter(Income < 17700)
ggplot(low_income_data, aes(x = Age_Bracket, fill = Age_Bracket)) +
  geom_bar() +
  geom_text(stat = "count", aes(label = ..count..), vjust = -0.5) +
  ggtitle("Number of People Earning Less Than £17,700 by Age Bracket") +
  xlab("Age Bracket") +
  ylab("Count") +
  scale_fill_manual(values = c("0-18" = "lightblue", "19-30" = "lightblue", "31-40" = "lightbl
  theme_minimal()
```

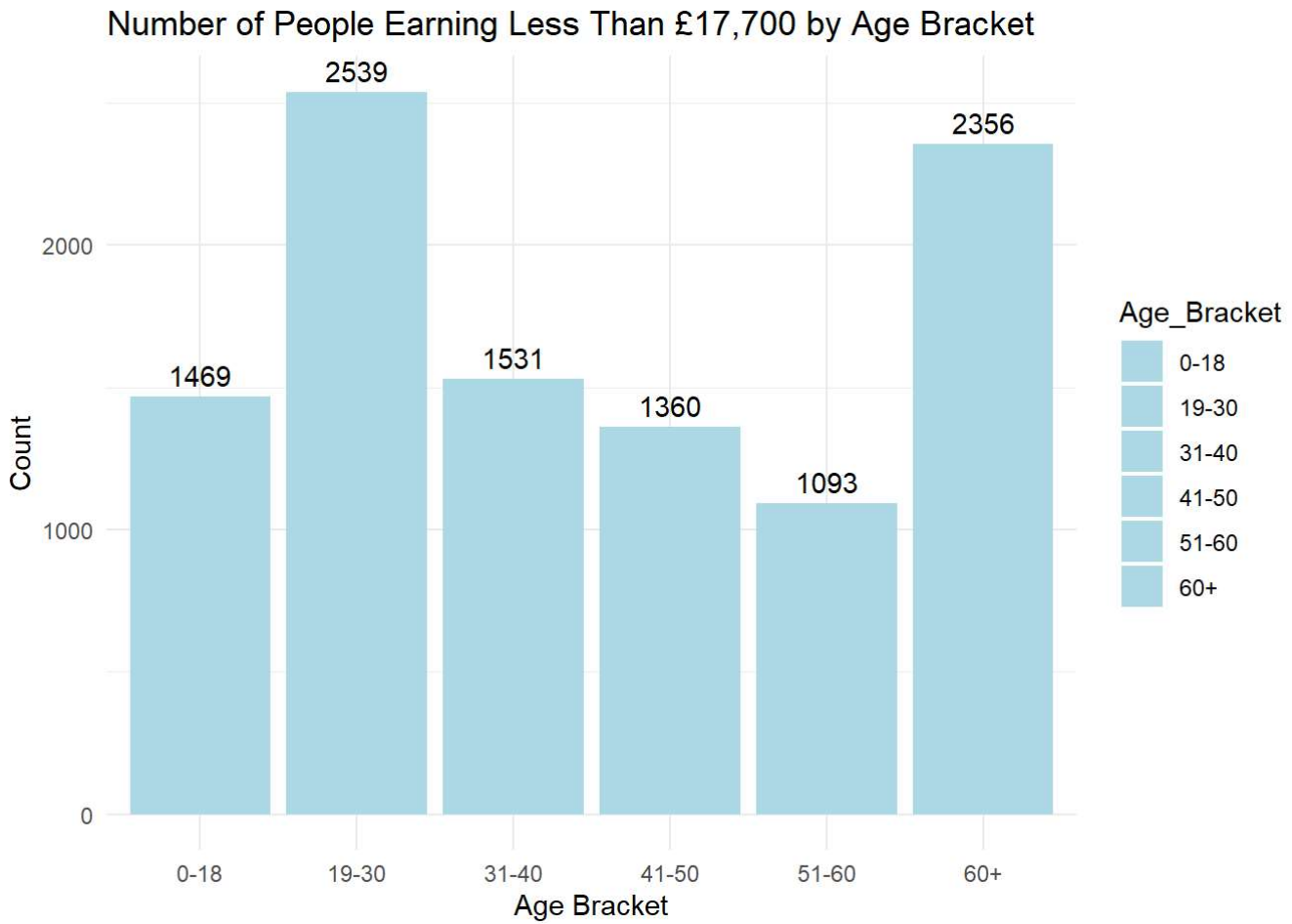


Figure 1: Count of People earning less than £17,700 by Age Group

Figure 1 shows the number of people in each age bracket who are earning less than 24,000. This data can help the government to bring the low income people up and make policies to give more opportunities to help each age bracket. From this results, it is clearly visible people age between 19-30 which is mostly youth are earning very less. After that comes the people age 60+ these are the mostly pensioners who are earning very less due to their age or health conditions. This 2827 count will help the government to make more pensioner schemes.

3.2 Effect of Level of Education attained on Low Income

▼ Code

```
low_income_data <- data |>
  filter(Income < 17700)
education_percentages <- low_income_data |>
  group_by(Highest_Education) |>
  summarise(Count = n()) |>
  mutate(Percentage = (Count / sum(Count)) * 100)
ggplot(education_percentages, aes(x = "", y = Percentage, fill = factor(Highest_Education))) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar(theta = "y") +
  ggtitle("Education Levels of People Earning Less Than £17,700 (Percentage)") +
  theme_void() +
  theme(legend.position = "right") +
  geom_text(aes(label = sprintf("%.1f%%", Percentage)), position = position_stack(vjust = 0.5)
```

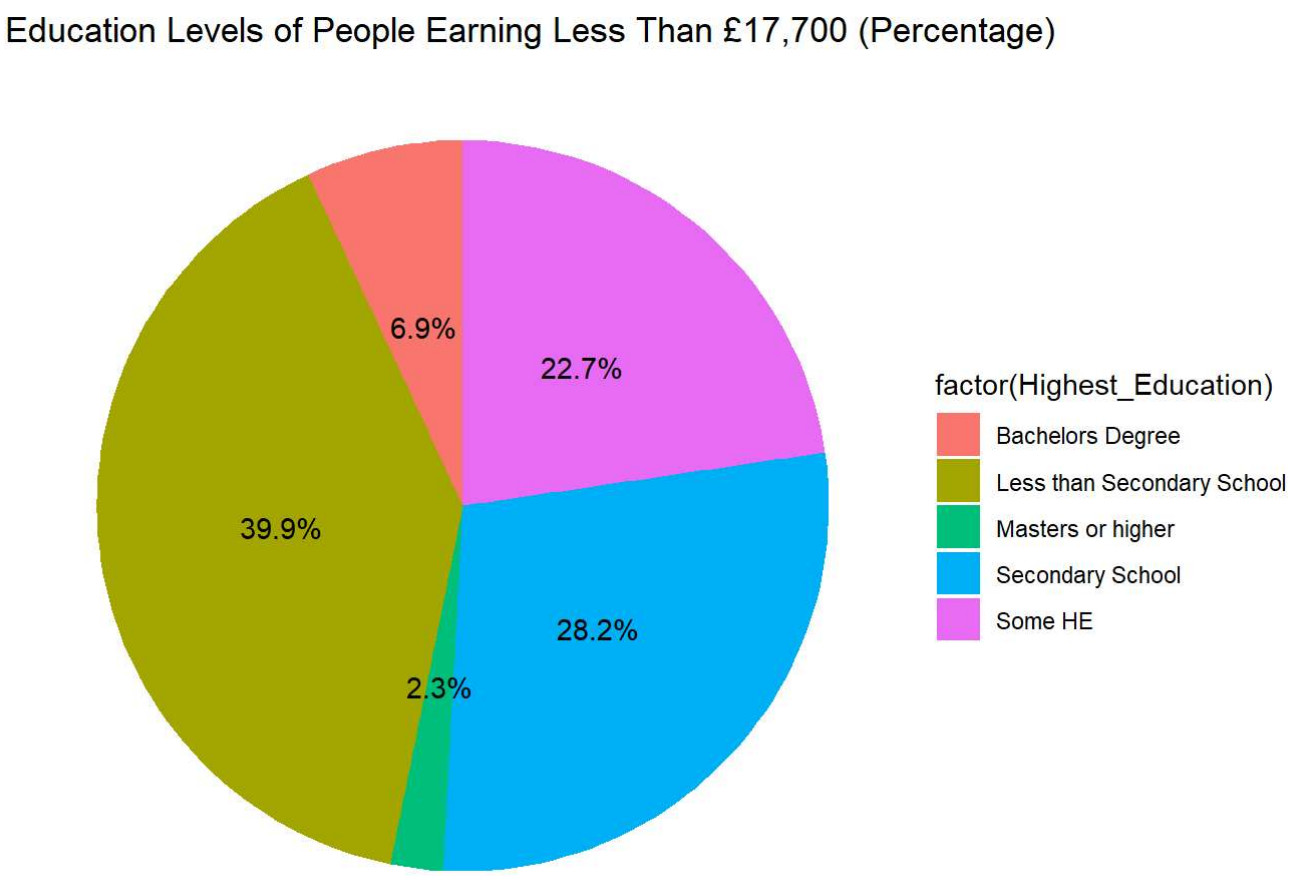


Figure 2: Education Level of People Earning less than £17,700

Figure 2 shows the one of the reason behind the low income individuals. It might be because of their education level. As the pie chart above shows that 36.3% people attended less than secondary school and went for low paying jobs or might get no opportunity for the jobs. People with bachelors degree or higher are much likely to earn more than £24,000. As, employers are more likely to give high paying jobs to people with some skills and advanced knowledge. In order to overcome this problem, Government should promote the higher education by providing the further assistance to students who attend university. Moreover they also expand the apprenticeships programs and make them more accessible to people of all ages and backgrounds.

3.3 Age Distribution of Married Inviduals by Ethnic Group

▼ Code

```
data |>
  filter(Marital_Status == "married") |>
ggplot(aes(x = Ethnicity, y = Age, fill = Ethnicity)) +
    geom_boxplot() +
    ggtitle("Age Distribution of Married Individuals by Ethnic Group") +
    xlab("Ethnic Group") +
    ylab("Age") +
    theme_minimal()
```

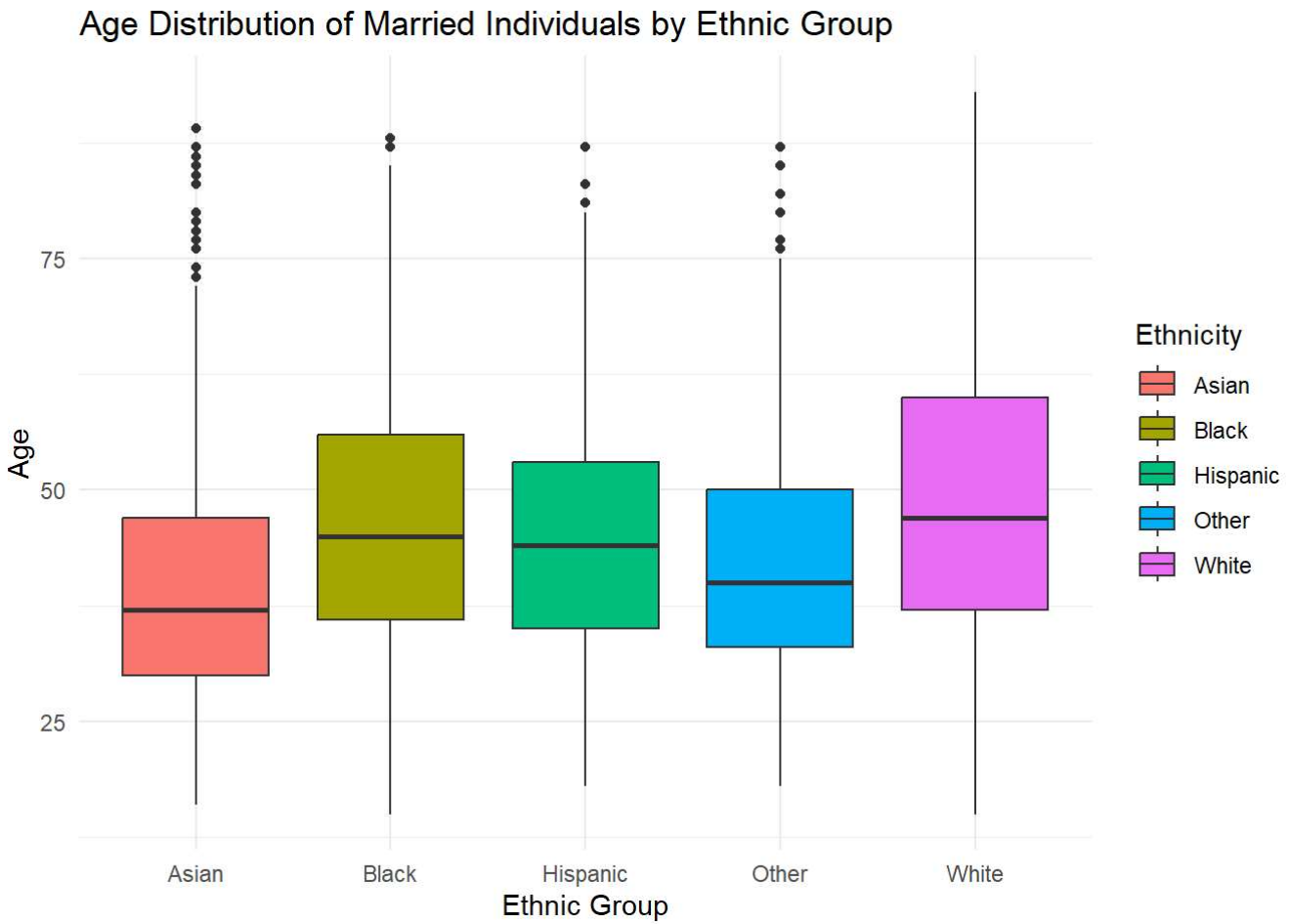


Figure 3: Age Distribution of Married Indivduals by Ethnic Group

Although, from Figure 3 Asian community seems less educated and earning less income, still the median age of married people in the asian community is lower than the median age of married people in any other asian community. Moreover the Figure 1 reflects that young people likely to be in low income category, despite the fact there seems to be more young married people in asian ethnicity. However, the age distribution says that median age of white people is around 45. From Figure 1 it has also been analysed that people age between 41-50 are earning more than any other age bracket. It concludes that married white people are more likely stable than married asian people.

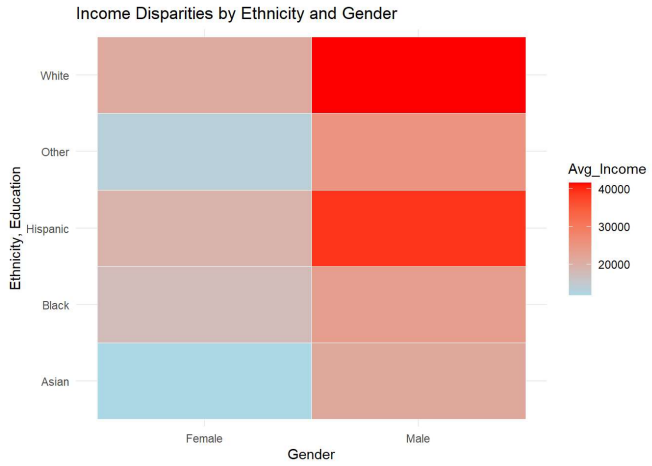
3.4 Income Variation by Ethnicity, Gender & Accomodation type of each Ethnic Group

▼ Code

```
income_summary <- data |>
   group_by(Ethnicity, Gender) |>
   summarise(Avg_Income = mean(Income, na.rm = TRUE))

ggplot(income_summary, aes(x = Gender, y = Ethnicity, fill = Avg_Income)) +
       geom_tile(color = "white") +
      ggtitle("Income Disparities by Ethnicity and Gender") +
      #lab("Gender (0 = Male, 1 = Female)") +
      ylab("Ethnicity, Education") +
      scale_fill_gradient(low = "lightblue", high = "red") +
      scale_color_colorblind() +
      theme_minimal()

ggplot(data, aes(Ethnicity)) +
      geom_bar(aes(fill=Accomodation_Type), width = 0.5) +
      theme(axis.text.x = element_text(angle=65, vjust=0.6)) +
      labs(title="People Living in Shared Accomodation by Ethnicity")
```



(a) Income Variation by Ethnicity and Gender

(b) Accomodation Type by Ethnicity

Figure 4: Plots

Figure 4 reflects the average income of each ethnicity and gender. It shows that **White and Hispanic male** are earning more than any other ethnicity. However, female in every ethnicity are like to earn less than male. To overcome this discrimination, Government should emphasize the employers to do gender balance in their organizations and make opportunities for females by introducing female oriented job schemes.
Moreover, the Figure 4 reflects that Asian community are earning very low (less than 20,000) it might be the reason only this ethnicity living in shared accommodation because they can't afford the self contained properties.

3.5 Count of Households with Individuals Aged over 60 by Marital Status

▼ Code

```
households_over_60 <- data %>%
     filter(Age > 60) %>%
```

```
    group_by(ID, Marital_Status) %>%
    summarise(Frequency = n())
ggplot(households_over_60, aes(x = Marital_Status, fill = Marital_Status)) +
     geom_bar() +
     geom_text(stat = "count", aes(label = ..count..), vjust = -0.5) +
     ggtitle("Count of Households with Individuals Aged Over 60 by Marital Status") +
     xlab("Marital Status") +
     ylab("Number of Households") +
     theme_minimal()
```

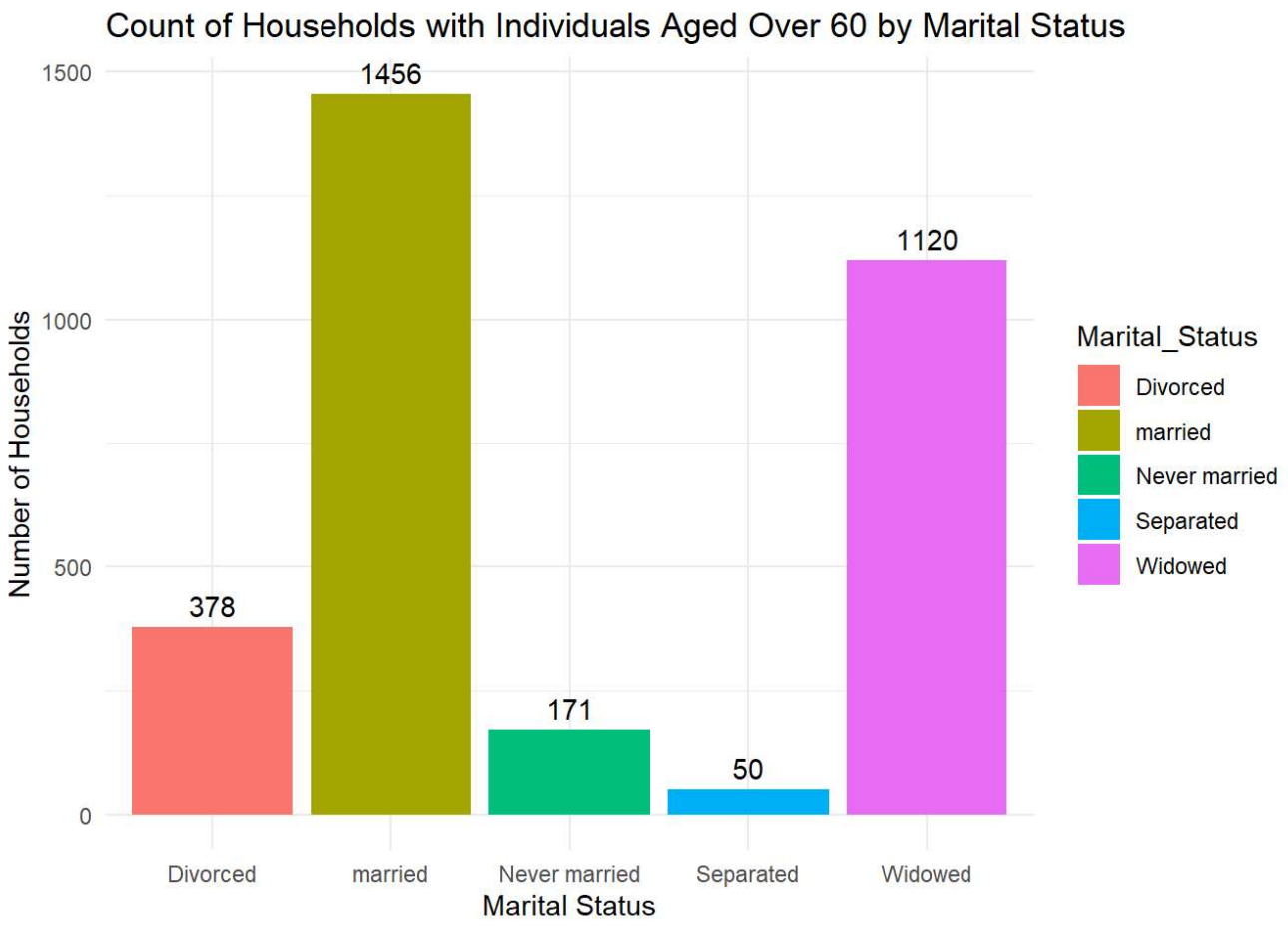


Figure 5: Count of Households with people age over 60 and their marital status

From the above Figure 5 , the number of households has been calculated where people is over 60 and their marital status. It shows that 1456 households has married old age people. there are few number of households with separated , never married and divorced people. This graph could help Government to analyse which households or number of households need career assistant. So old aged people who are living alone might get help from this data.

4 Conclusion

This analysis provided insights into the demographic and socioeconomic traits of people in the UK in 2021. It highlighted several key findings, including:

- A significant portion of the population (63.7%) has an annual income less than £17,700.
- Low-income individuals are more likely to be young (19-30) or old (60+).
- Educational attainment is a significant factor in determining income levels, with those with less than a secondary education being more likely to be in low-income households.
- Income disparities exist between ethnicities and genders, with White and Hispanic males earning more than other groups.
- Asian individuals are more likely to live in shared accommodation, potentially due to lower income levels.
- A considerable number of households (1456) have individuals aged over 60 who are married.

These findings suggest that there is a need for policies and interventions to address low income, educational disparities, and income inequalities. Additionally, support systems for older individuals living alone should be considered. By addressing these issues, the UK can work towards a more equitable and just society for all.

References

GOV.UK. 2023. “People in low income households - GOV.UK Ethnicity facts and figures.” https://www.ethnicity-facts-figures.service.gov.uk/work-pay-and-benefits/pay-and-income/low-income/latest.

National Statistics, Office for. 2023. “Office for National Statistics.” https://www.ons.gov.uk/.

Wickham, Hadley. 2016. “Ggplot2: Elegant Graphics for Data Analysis.” https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the Tidyverse” 4: 1686. https://doi.org/10.21105/joss.01686.