TO MEAT, OR NOT TO MEAT-

Can Vegetarianism Boost Your Happiness? Analyzing the Relationship between an Individual's Diet Preference and Life Satisfaction.

Litta Jose Thottam, 135546

2024-08-30

Contents

1 Introduction **2**

2 Main **5**

2.1 Data 5

2.2 Variables 6

2.3 Data Management 7

2.4 Sample 7

3 Data Analysis and Data Visualizations(Variables) **7**

3.1 Multi Variable Analysis/Group Analysis 7

3.2 Individual Analysis 13

3.3 Vegetarinism and Current-life satisfaction(Happiness) 21

4 Overall Life satisfaction and Current Life satisfaction-assessment **24**

5 Regression and Co-relation. **25**

6 Presence of Outliers-Summary **26**

7 Limitations and Next Steps for In-Depth Analysis **26**

7.1 Limitations 26

7.2 Next Steps for In-Depth Analysis: 27

8 Conclusion **27**

References **28**

1 Introduction

Human acquires special behaviors under the effect of different internal and environmental factors and accordingly has a unique personality. For this reason, one of the main issues in psychology is individual's personality traits.(Aslanifar et al., 2014)

In an era where Veganism and Vegetarianism are hot topics on morality and ethics ,it is important to delve into another key aspect-of happiness/life satisfaction. The number of vegetarians worldwide is estimated to be 1 billion.In Germany the number is rising quickly and the number of vegans even more so. At the moment there are 800,000 vegans in Germany and the number has been rising exponentially.

The studies referenced in this project highlight the nuanced relationship between dietary choices, specifically vegetarianism, and individual happiness. While traditional health and wellness theories suggest that diet can have a significant impact on mental well-being, the findings reveal a more complex interplay influenced by cultural norms, social environment, and personal beliefs.

A study in 2010 explores the mental health benefits of vegetarian diets12. Conducted by Beezhold et al., the research involved 138 healthy Seventh Day Adventist men and women. The study found that vegetarians reported significantly less negative emotion compared to omnivores, despite lower intakes of long-chain omega-3 fatty acids like EPA and DHA3. The results suggest that vegetarian diets, which are higher in polyunsaturated fats and lower in arachidonic acid, may contribute to better mood states4. This challenges the notion that low omega-3 intake adversely affects mental health and highlights the potential mental health benefits of vegetarian diets5.. However, other studies challenge this notion, indicating that the relationship is not straightforward and can vary widely depending on factors such as social support, cultural acceptance, and access to quality plant-based foods.(Beezhold et al., 2010)

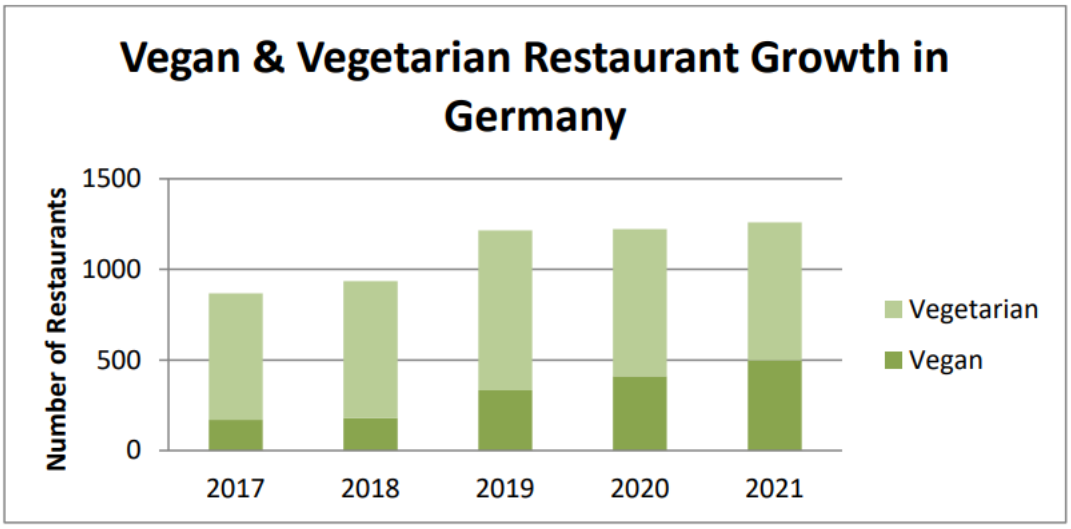
Interestingly, the role of social identity in dietary choices is a recurring theme, with vegetarians often reporting a sense of belonging to a like-minded community, which can enhance their overall well-being.

A study in 2018 aims to estimate the prevalence of vegetarians, analyze socio-demographic influences on dietary behavior, and examine personality differences between vegetarians and meat eaters1. Findings indicate that vegetarians are more likely to be female, younger, and more educated. The study concludes that individual differences in socio-demographics, personality traits, and political attitudes significantly influence dietary choices.(Pfeiler & Egloff, 2018)

When examining the relationship through the lens of socio-demographic variables, this study found that age, gender, and income can moderate the impact of vegetarianism on happiness. For instance, younger individuals and women are more likely to report higher happiness levels associated with vegetarianism, possibly due to greater social support and alignment with contemporary values. This complex relationship underscores the need to consider a range of factors when evaluating the impact of dietary choices on happiness, moving beyond simplistic associations to understand the broader context.

The findings have important implications for public health initiatives and dietary recommendations, **suggesting that promoting vegetarianism could enhance well-being, but only when supported by a conducive social environment and accessible food options.** Moreover, the research highlights the importance of individual choice and the personal meaning attached to dietary practices, which can significantly influence their impact on happiness.

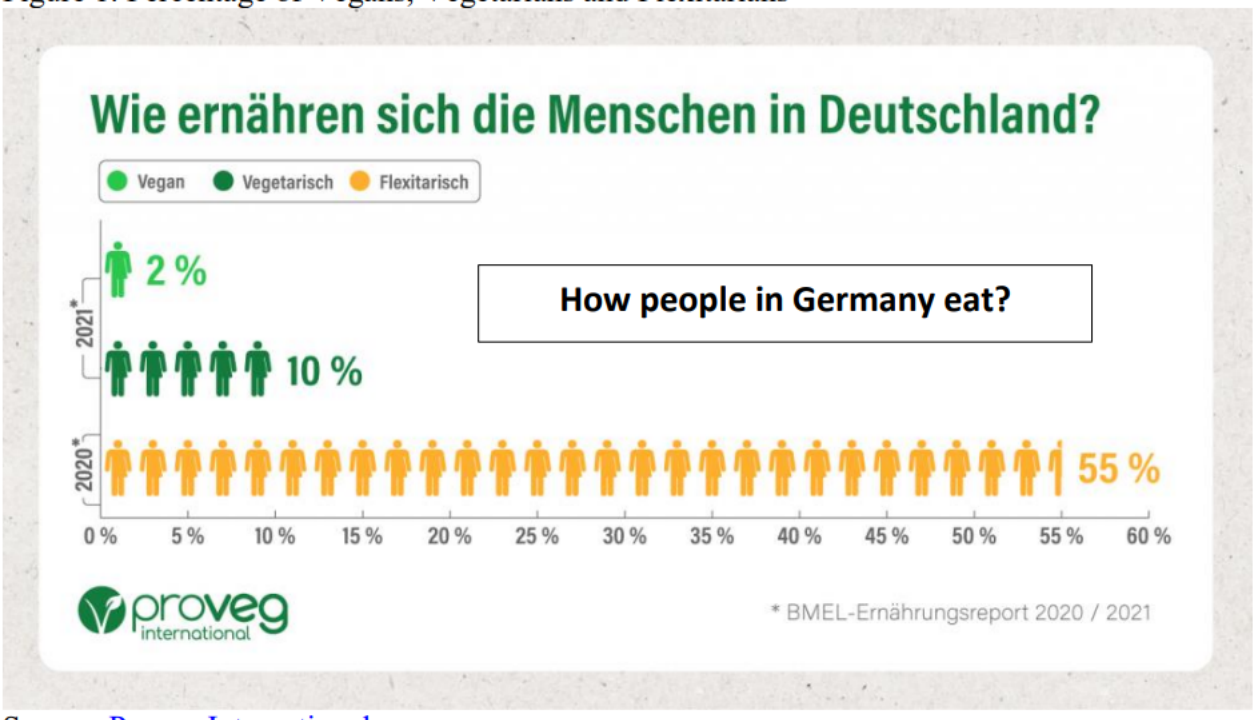
```
knitr::include_graphics("S2.png")
```



Source: HappyCow

```
knitr::include_graphics("S3.png")
```

Figure 1: Percentage of Vegans, Vegetarians and Flexitarians



Source: Proveg International

From these two figures its evident that the number of vegeteraians and vegans are increasing ,especially in germany and germany is second in the largest number of plant based people in Europe.

In recent years, the vegan and vegetarian population in Germany has rapidly expanded. According to a recent survey, nearly 8 million people followed a vegetarian diet and 1.58 million people identified themselves as vegan in 2022. In other words, nearly 10 million people are choosing to follow a diet without meat and fish or entirely without animal products. This is quite a change, since it is worth remembering that it was

0.1 million people 10 years ago who considered themselves vegan.Additionally, the number of people who are concerned with their meat consumption continues to grow:55% of people in Germany are "part-time vegetarians" or "flexitarians". Flexitarians often reduce but don't stop eating meat.(Rehder, 2023)

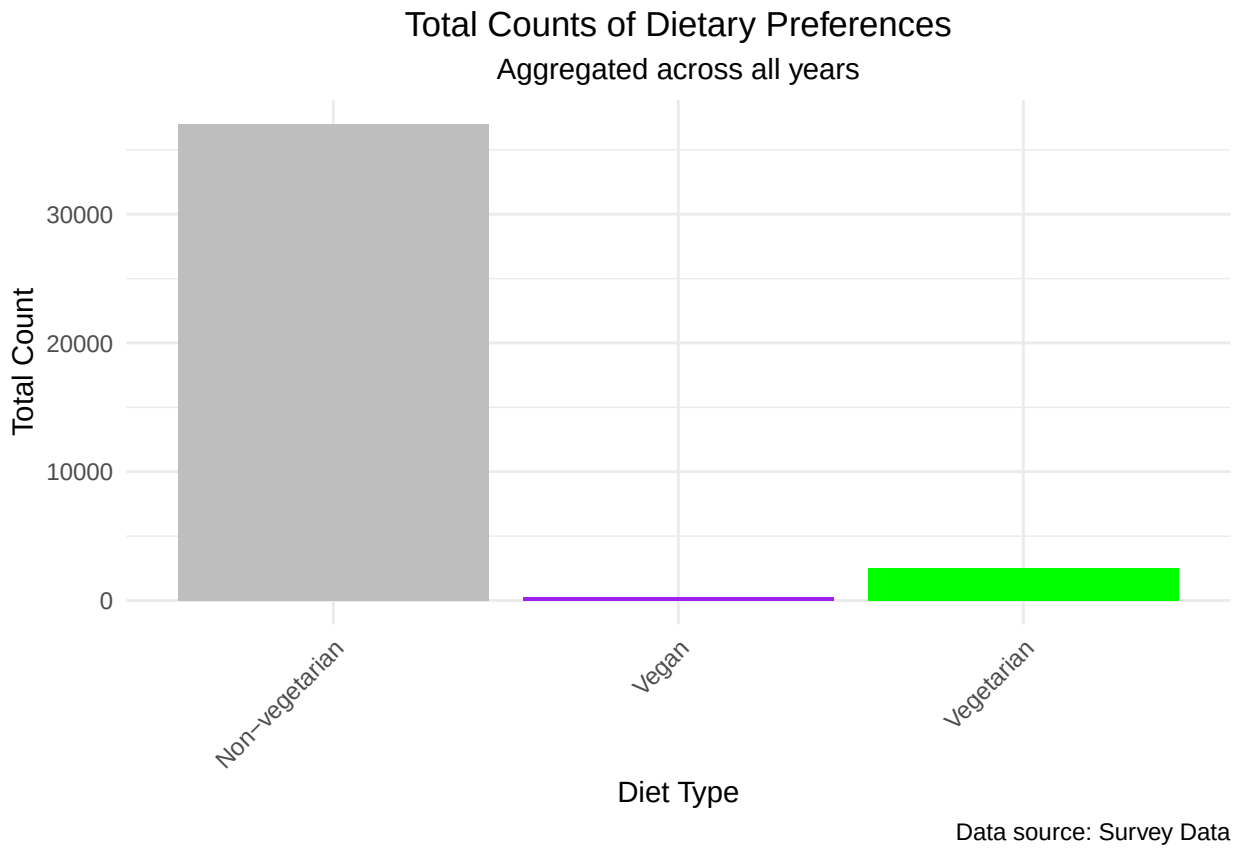
The German Socio-Economic Panel (SOEP) dataset, renowned for its longitudinal nature and wealth of variables, serves as the foundation for my project. A preliminary analysis of the data reveals intriguing patterns. A visual representation of the density plot allowing you to see how the average commuting times vary across different individuals.

```
  # Assuming you might need to read the data from a file

# If 'pl' is not already loaded, load it
# pl <- read_dta("pl.dta", col_select = c("pid", "syear", "ple0182"))

# Transform 'pl' to include DietType
pl <- pl %>%
  mutate(DietType = case_when(
    ple0182 == 1 ~ "Vegetarian",
    ple0182 == 2 ~ "Vegan",
    ple0182 == 3 ~ "Non-vegetarian"
  ))

# Filter out only relevant dietary types if necessary
pl_filtered <- pl[pl$ple0182 %in% c(1, 2, 3), ]

# Group by year and DietType, then summarize to count the number of each type per year
yearly_counts <- pl_filtered %>%
  group_by(syear, DietType) %>%
  summarise(Count = n(), .groups = 'drop')

# Calculate the average number of each DietType across all years
average_counts <- yearly_counts %>%
  group_by(DietType) %>%
  summarise(AverageCount = mean(Count))

# Print out the results
print(average_counts)
```

```
## # A tibble: 3 x 2
##   DietType       AverageCount
##   <chr>                 <dbl>
## 1 Non-vegetarian       12340.
## 2 Vegan                   99
## 3 Vegetarian             852
```

```
total_counts <- pl_filtered %>%
  group_by(DietType) %>%
  summarise(TotalCount = n(), .groups = 'drop')

# Plotting the histogram
ggplot(total_counts, aes(x = DietType, y = TotalCount, fill = DietType)) +
  geom_bar(stat = "identity") +  # Using geom_bar which requires stat = "identity" to use y values
  scale_fill_manual(values = c("Vegetarian" = "green", "Vegan" = "purple", "Non-vegetarian" = "gray")) +
  labs(x = "Diet Type", y = "Total Count",
```

```
      title = "Total Counts of Dietary Preferences",
      subtitle = "Aggregated across all years",
      caption = "Data source: Survey Data") +
 theme_minimal() +
 theme(plot.title = element_text(hjust = 0.5),
       plot.subtitle = element_text(hjust = 0.5),
       plot.caption = element_text(hjust = 1),
       legend.position = "none",  # Remove legend if unnecessary
       axis.text.x = element_text(angle = 45, hjust = 1))
```



The primary aim of this project is to explore whether there is a discernible correlation between vegetarianism and levels of happiness among individuals. Are individuals who follow a vegetarian diet more likely to report higher levels of happiness? In an era of increasing awareness about health and environmental impacts of dietary choices, such insights hold the potential to inform policies that promote not only health and well-being, but also sustainable practices that contribute to the welfare of the environment and society at large.

2 Main

2.1 Data

SOEP is a comprehensive and longitudinal household survey conducted in Germany. Initiated in 1984, the primary purpose of SOEP is to provide insights into the dynamics of social and economic conditions, labor market behavior, and overall well-being among German households. Its scope encompasses a wide range

of topics, offering a deep understanding of individual and household characteristics, preferences, and trends (Goebel et al., 2019).

The data collection methodology of SOEP involves face-to-face interviews conducted annually. This approach allows for the collection of detailed and accurate information directly from respondents. The survey includes a diverse set of households, covering various demographic groups, socio-economic backgrounds, and geographic regions. This sampling approach ensures that the dataset is representative of the broader German population.

The study's sample size is substantial, consisting of around 20,000 private households and over 30,000 individuals each year. The sample composition is carefully selected to ensure representation across different age groups, income levels, educational backgrounds, and household structures. This diversity enables researchers to analyze changes and trends over time within specific demographic categories, enhancing the study's analytical capabilities.

SOEP covers an extensive array of key variables, ranging from demographic details and labor market characteristics to income, health, education, and housing. The availability of such comprehensive data makes SOEP a valuable resource for researchers across various disciplines, facilitating in-depth analyses of complex social and economic phenomena.

A distinctive feature of SOEP is its longitudinal nature. The study's data collection occurs annually, allowing researchers to examine changes, trajectories, and patterns over an extended period. This temporal dimension enables the investigation of life course dynamics, intergenerational trends, and the effects of policy changes, fostering a deep understanding of social and economic shifts.

Access to SOEP data is granted to researchers upon application and approval, subject to certain usage restrictions and confidentiality measures. To protect respondents' privacy, the data undergoes rigorous anonymization procedures, ensuring that no individual can be identified from the dataset. Researchers are bound by ethical guidelines to handle the data with care and to maintain confidentiality.

2.2 Variables

The main variables selected for this project are vegetarinism and Current Life Satisfaction. These were taken from the pl dataset. It contains individual-level data for respondents by answering the annual individual questionnaire. It is keyed on PID (Person ID) and SYEAR (Survey Year).

Vegitarinism /Vegan: Vegetarinism/Vegane empfahnung refers whwther the person is under a plant based Diet or not.

Current Life Satisfaction: Current life satisfacton is a good indicator for the current state of mind (happiness or else) of an individual at any given point of time.

Relationship between Vegeterinism and Happiness: The relationship between vegeterinism and happiness examines how these two variables are interrelated within the context of individuals' daily lives and happiness.

Other Socio-Economic Factors:

1.Age Age of the individual, which may influence both dietary choices and life satisfaction.

2.Gender Gender of the individual, as it could affect dietary preferences.

3 Income Of Individual The individual's income level, which may impact both their ability to follow a vegetarian diet and their overall life satisfaction.

4 Education The highest level of education attained, which might correlate with dietary choices

5 Health Self-reported health status, as it may influence the decision to follow a vegetarian diet.

6 Survey Year The year of the survey, to account for temporal changes in dietary trends.

7 Relegion Religious affiliation, which can affect dietary practices,for example ,omission of certain food products or strictly plant based diets(Hindus,Buddhists,Jains).

In Addition,anoher happiness indicator-**Overall Life satisfaction** is also taken and analysed inorder to find if the current life satisfaction differs highly.If Overall Life satisfaction gives almost the same trend or result,we can make sure current life satisfaction is reliable.

In total, **10 Variables** have been selected and analyzed for this project. These variables were chosen after exploring more than **15 potential variables** from the SOEP dataset. The selected variables provide a comprehensive overview of the factors influencing the relationship between vegetarianism and happiness, including socio-demographic characteristics, income, health, education, and religious affiliation.

When considering the relationship between happiness and Vegetarianism,which is our main topic,we explore whether there exists a correlation between an individuals dietary preference and happiness. In other words, we investigate whether people who eat a plant based diet are more happy or not compared to meat eaters.

2.3 Data Management

These are the SOEP variable names used in the project.

- **ple0182** Vegetarianism - Represents the individual's dietary preference categorized as Vegetarian, Vegan, or Meat Eater.
- **plh0182**: Current Life Satisfaction - The individual's self-reported level of happiness or life satisfaction.
- **d11101**: Age - The age of the individual.
- **d11102ll**: Gender - The gender of the individual, categorized as Man or Woman.
- **pglabnet**: Net Labor Income - The individual's reported net income from labor after deductions.
- **m11126**: Self-Rated Health - The individual's self-assessed health status.
- **d11108**: Education Level - The highest level of education attained by the individual.
- **plh0258_h**: Religion - The individual's religious affiliation, categorized into different religious groups.
- **syear**: Survey Year - The year in which the survey was conducted.
- **p11101**: Overall Life Satisfaction - An additional measure of the individual's overall life satisfaction, used to cross-validate the current life satisfaction variable.

Total variables taken:10

2.4 Sample

The estimation sample is derived from the dataset pl.dta,pequiv.dta,pgen.dta after applying data preparation steps including excluding rows with certain exclusion values various columns that represent the needed variables. These excluded ones are data that are missing, lacking a valid code or value or for other different reasons.

3 Data Analysis and Data Visualizations(Variables)

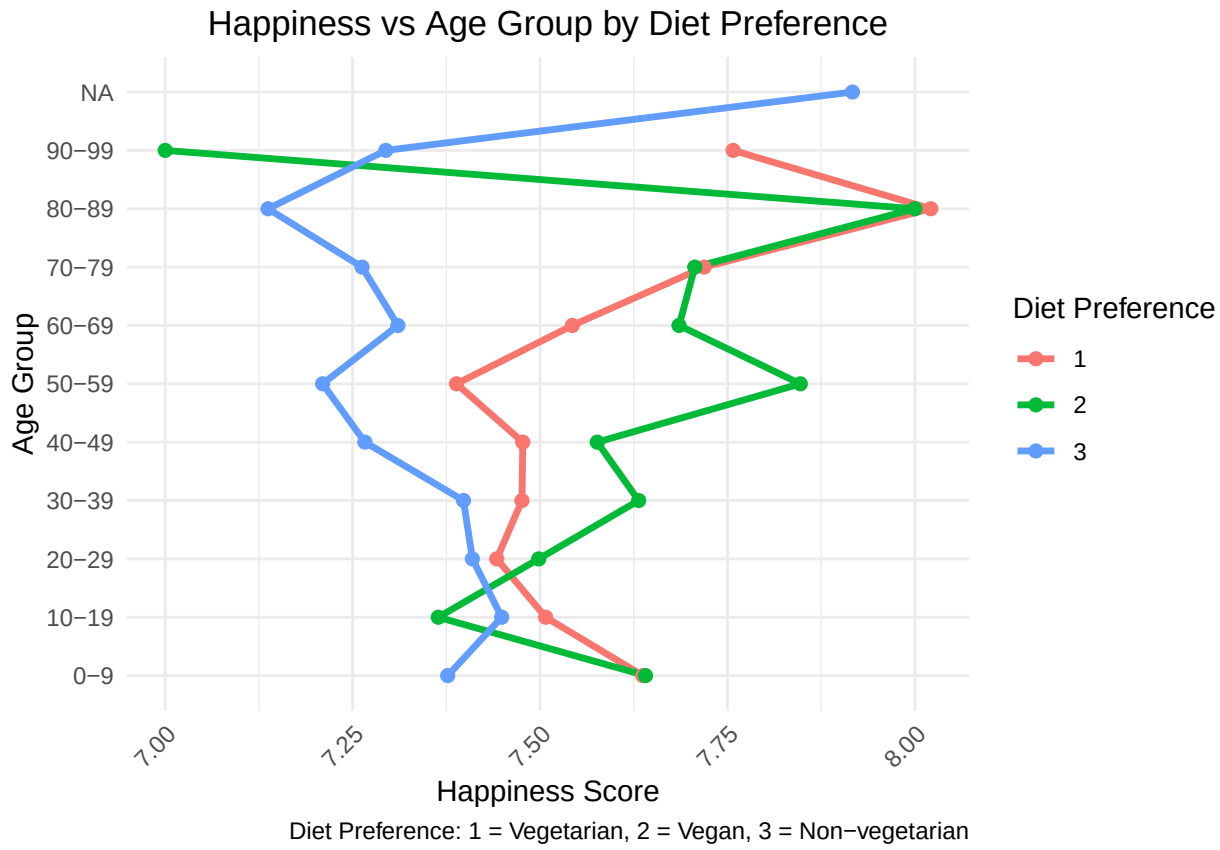
3.1 Multi Variable Analysis/Group Analysis

The following variables are analysed with 2 other variables(vegetarinism and current life satisfaction)

3.1.1 Age and Vegetarianism

With Age (d11101) from pequiv data,we analyse if there is any correlation between age and vegetarianism.

```
# Create age groups (optional)
merged_data <- merged_data %>%
  mutate(age_group = cut(age, breaks = seq(0, 100, by = 10), right = FALSE, labels = c("0-9", "10-19",

# Line plot with happiness on X-axis, age groups on Y-axis, and diet preferences as lines
ggplot(merged_data, aes(x = plh0182, y = age_group, color = as.factor(ple0182), group = as.factor(ple018
  geom_line(stat = "summary", fun = mean, size = 1.2) +
  geom_point(stat = "summary", fun = mean, size = 2) +
  labs(title = "Happiness vs Age Group by Diet Preference",
       x = "Happiness Score",
       y = "Age Group",
       color = "Diet Preference",
       caption = "Diet Preference: 1 = Vegetarian, 2 = Vegan, 3 = Non-vegetarian") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5),
        axis.text.x = element_text(angle = 45, hjust = 1))
```



From the graph, it is evident that non-vegetarians (Diet Preference 3, blue line) generally report higher happiness scores across most age groups compared to vegetarians (Diet Preference 1, red line) and vegans (Diet Preference 2, green line). The pattern of happiness scores varies with age, with non-vegetarians showing relatively stable happiness levels across age groups, while vegetarians and vegans display more fluctuation, particularly in younger and middle-aged groups. Interestingly, there is a noticeable gap in happiness scores

between the different diet groups, especially in the 50-59 age group, where vegetarians and vegans show a decline, while non-vegetarians maintain higher levels of happiness.

Outliers: There don't appear to be significant outliers in this graph, but there is a noticeable dip in happiness for vegetarians and vegans in the 50-59 age group, which could be considered an outlier if it contrasts sharply with trends in adjacent age groups.

3.1.2 Gender And Vegetarinism

```
# Plot happiness scores by diet preference, faceted by gender
ggplot(merged_data, aes(x = as.factor(ple0182), y = plh0182, fill = as.factor(ple0182))) +
  geom_boxplot() +
  facet_wrap(~ gender, scales = "free") +
  labs(
    title = "Happiness Scores by Diet Preference and Gender",
    x = "Diet Preference",
    y = "Happiness Score",
    fill = "Diet Preference"
  ) +
  scale_fill_brewer(palette = "Pastel1") +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5),
    axis.text.x = element_text(angle = 45, hjust = 1),
    legend.position = "bottom"
  )
```

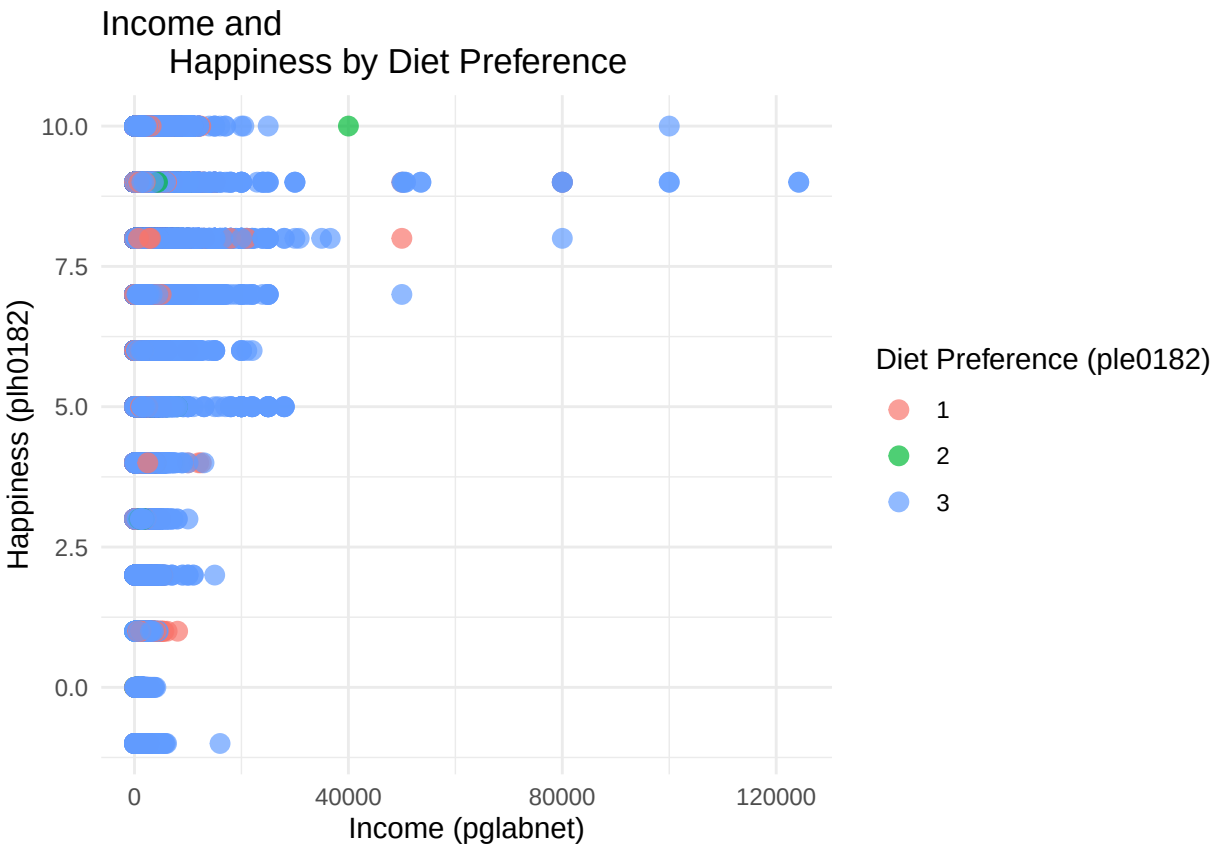
From the graph, it is evident that vegetarians (Diet Preference 1, red) generally report slightly higher happiness scores compared to vegans (Diet Preference 2, blue) and non-vegetarians (Diet Preference 3, green) across both men and women. Among men, vegetarians show a wider range of happiness scores, indicating more variability in their well-being, while non-vegetarians have more consistent happiness levels. For women, the happiness scores are more evenly distributed across all diet preferences, but vegetarians still appear to have a slight edge in median happiness. Interestingly, in the "NA" category,which could be taken as Non-Binary (where gender is unspecified), non-vegetarians report relatively high and consistent happiness scores, suggesting this group, whether non-binary or unspecified, experiences stable well-being. This consistency is particularly notable as it contrasts with the more variable scores observed in other groups.

Outliers: There are individual points (likely representing outliers) at the lower end of the happiness scale, particularly among men and women who are vegetarians. These indicate individuals reporting significantly lower happiness than the majority of their group.

3.1.3 Income and Vegetarinism

```r
# Load necessary libraries
library(haven)
library(dplyr)

# Load the data
pl_income <- read_dta("pl.dta", col_select = c("pid", "syear", "ple0182", "plh0182"))
pgen_income <- read_dta("pgen.dta", col_select = c("pid", "pglabnet"))

# Exclusion values
```

```r
exclusion_values <- c(-1, -2, -3, -4, -5, -6, -7, -8)

# Remove rows with exclusion values in the plb0585 column from pl_income
pl_income <- pl_income[!pl_income$ple0182 %in% exclusion_values, ]

# Remove rows with exclusion values in the pglabnet column from pgen_income
pgen_income <- pgen_income[!pgen_income$pglabnet %in% exclusion_values, ]

# Merge the pgen_income data with the filtered pl_income data
merged_data <- merge(pl_income, pgen_income, by = "pid", all.x = TRUE)



# Plot
ggplot(merged_data, aes(x = pglabnet, y = plh0182, color = as.factor(ple0182))) +
  geom_point(size = 3, alpha = 0.7) +
  labs(title = "Income and
       Happiness by Diet Preference",
       x = "Income (pglabnet)",
       y = "Happiness (plh0182)",
       color = "Diet Preference (ple0182)") +
  theme_minimal()
```



From the graph, it is evident that non-vegetarians (Diet Preference 3, blue) generally occupy higher income brackets and display a wide range of happiness scores, with a concentration around the higher end of the happiness spectrum.

This suggests that non-vegetarians with higher incomes tend to report greater levels of happiness. On the other hand, vegetarians (Diet Preference 1, red) and vegans (Diet Preference 2, green) are more dispersed across lower to middle income levels, with vegetarians showing slightly higher happiness scores compared to vegans. *Interestingly, vegans, despite often having lower income levels, sometimes report very high happiness scores, indicating that their well-being might be influenced by factors other than income.* Overall, non-vegetarians exhibit a more direct correlation between higher income and happiness, while vegetarians and vegans display more variability in their happiness scores, independent of income level.

Outliers: The presence of non-vegetarians with high income but low happiness could be considered outliers, as higher income typically correlates with higher happiness. Similarly, the very high happiness scores among vegans at lower income levels might be seen as outliers if they deviate from the expected trend.

3.1.4 Self-Health Rating and Vegitarinism

This chart tries to analyse whether an individual's diet preference make them think that they are healthier,Assuming that people who consider themselves to be healthier would be happier.

```
# Load the data
pl_health <- read_dta("pl.dta", col_select = c("pid", "syear", "ple0182", "plh0182"))
pequiv_health <- read_dta("pequiv.dta", col_select = c("pid", "m11126"))

# Exclusion values
exclusion_values <- c(-1, -2, -3, -4, -5, -6, -7, -8)

pl_health <- pl_health[!pl_health$ple0182 %in% exclusion_values, ]

# Remove rows with exclusion values in the m11126 column from pgen_income
pequiv_health <- pequiv_health[!pequiv_health$m11126 %in% exclusion_values, ]

# Merge the health
merged_health <- merge(pl_health, pequiv_health, by = "pid", all.x = TRUE)


# Plot
ggplot(merged_health, aes(x = m11126, y = plh0182, color = as.factor(ple0182))) +
  geom_point(size = 3, alpha = 0.7) +
  labs(title = "Self Rated  Health by Diet Preference",
       x = "Self rated health (m11126)",
       y = "Happiness (plh0182)",
       color = "Diet Preference (ple0182)") +
  theme_minimal()
```

From the graph, it is evident that non-vegetarians (Diet Preference 3, blue) generally dominate across all levels of self-rated health, with a significant presence at both lower and higher happiness scores. This suggests that non-vegetarians, regardless of how they rate their own health, experience a wide range of happiness levels, indicating that factors other than self-perceived health might be influencing their overall happiness.

Vegetarians (Diet Preference 1, red) and vegans (Diet Preference 2, green) are more sparsely distributed across the health ratings. Vegetarians, in particular, seem to cluster more in the mid to higher self-rated health categories (3 to 5), often reporting higher happiness scores as their self-rated health improves. This suggests that vegetarians who consider themselves healthier are more likely to report greater happiness.

Interestingly, vegans, while fewer in number, occasionally report very high happiness scores even at mid-levels of self-rated health. This indicates that vegans' happiness might be less directly tied to their self-rated health compared to vegetarians and non-vegetarians, potentially reflecting other contributing factors to their well-being.

Outliers: There are some low happiness scores among individuals with high self-rated health, particularly among non-vegetarians, which could be seen as outliers since better health typically correlates with higher happiness.

3.2 Individual Analysis

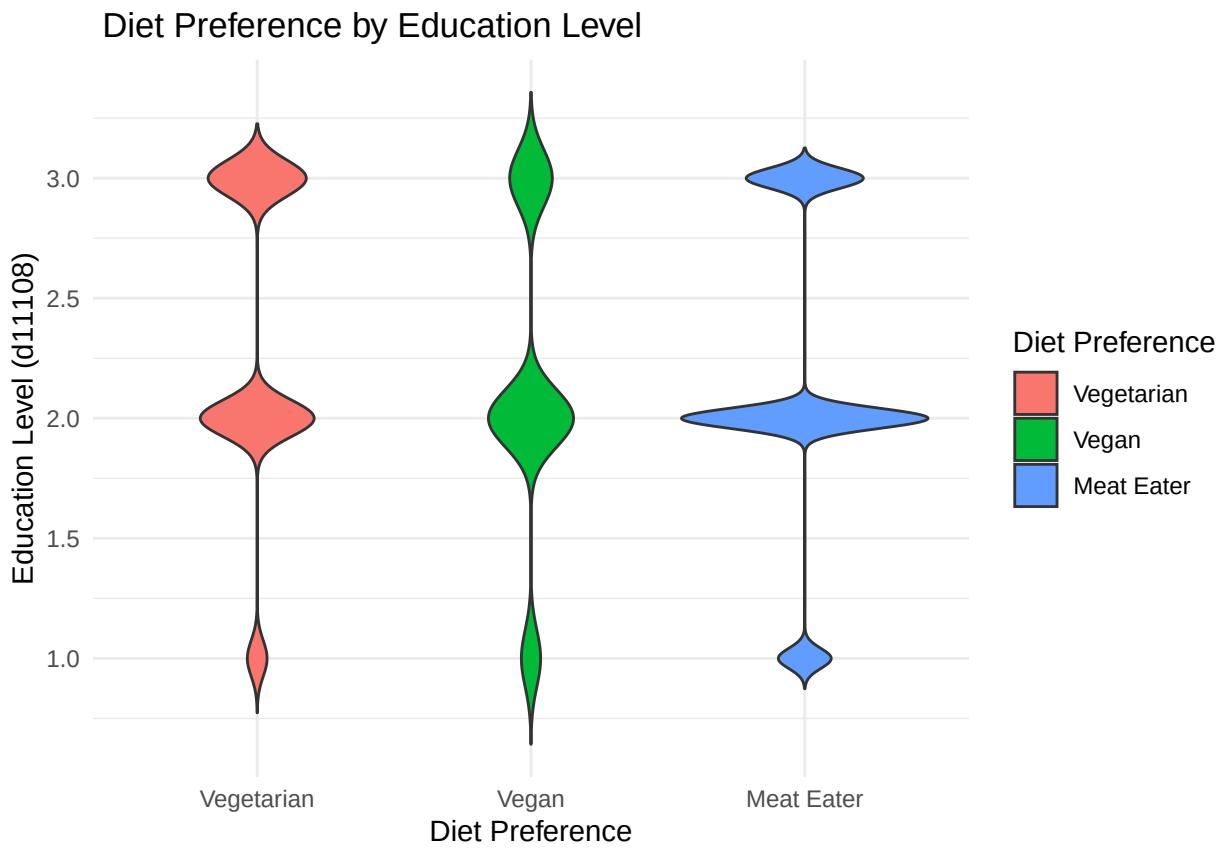
The following graphs are plotted and analysed just with the variable and dietary preference.

3.2.1 Education and vegetarianism

This analysis aims to explore the relationship between an individual's education level and their reported happiness, under the assumption that higher education levels might correlate with greater happiness. The

rationale behind this assumption is that higher education often provides individuals with better opportunities, higher income potential, and greater personal fulfillment, all of which could contribute to increased life satisfaction and happiness.

```
# Load the data
pl_education <- read_dta("pl.dta", col_select = c("pid", "syear", "ple0182", "plh0182"))
pequiv_education <- read_dta("pequiv.dta", col_select = c("pid", "d11108"))

# Exclusion values
exclusion_values <- c(-1, -2, -3, -4, -5, -6, -7, -8)

pl_education<- pl_education[!pl_education$ple0182 %in% exclusion_values, ]

# Remove rows with exclusion values in the m11126 column from pgen_income
pequiv_education <- pequiv_education[!pequiv_education$d11108 %in% exclusion_values, ]

# Merge the education
merged_education <- merge(pl_health, pequiv_education, by = "pid", all.x = TRUE)



merged_education$ple0182 <- factor(merged_education$ple0182,
                                   levels = c(1, 2, 3),
                                   labels = c("Vegetarian", "Vegan", "Meat Eater"))

# Violin plot
ggplot(merged_education, aes(x = ple0182, y = d11108, fill = ple0182)) +
  geom_violin(trim = FALSE) +
  labs(title = " Diet Preference by Education Level ",
       x = "Diet Preference",
       y = "Education Level (d11108)",
       fill = "Diet Preference") +
  theme_minimal()
```



This graph illustrates the distribution of education levels across different diet preferences, specifically for vegetarians, vegans, and meat eaters. The x-axis represents the three diet categories, while the y-axis shows the education levels. Each violin plot displays the spread and density of individuals within each diet group according to their education level.

From the graph, it is evident that vegans (green) generally have higher education levels compared to vegetarians (red) and meat eaters (blue). Vegans show a strong concentration at the upper end of the education spectrum, particularly around level 3, indicating that a significant portion of vegans have attained higher education. Vegetarians also tend to have higher education levels, though their distribution shows a more balanced presence between levels 2 and 3, with fewer individuals at the lowest education level.

Meat eaters, on the other hand, exhibit a more even distribution across the education spectrum, with a noticeable concentration around level 2. This suggests that meat eaters come from a broader range of educational backgrounds, with representation across both lower and higher education levels.

The shape and spread of these violin plots suggest that higher education levels are more closely associated with vegetarian and vegan diets, while meat eaters tend to have a more diverse range of educational backgrounds. This distribution implies that educational attainment may influence dietary choices, with individuals who have higher education being more likely to adopt vegetarian or vegan diets.

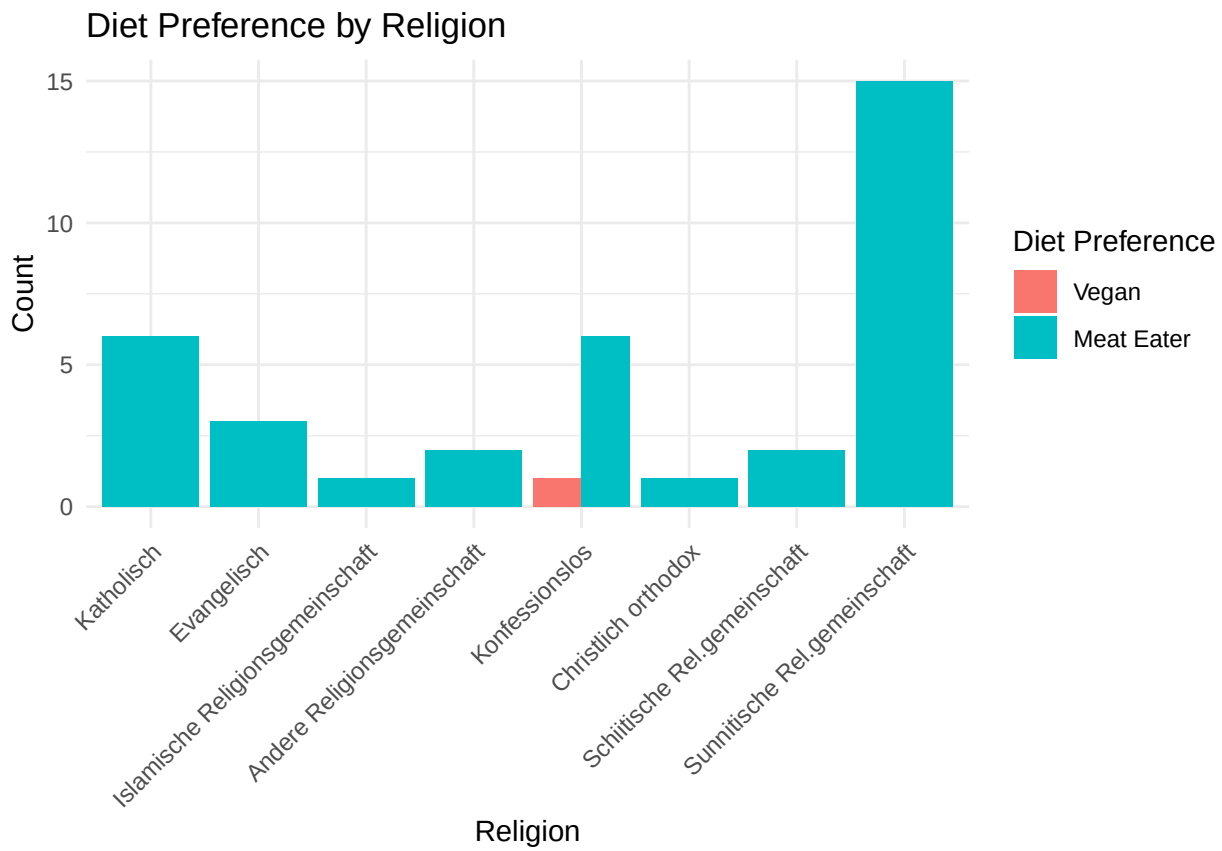
Outliers: The plot suggests a concentration at the extremes (e.g., vegetarians at the highest education level), but these are not necessarily outliers—they represent the distribution within each group. 7. Diet Preferen

3.2.2 Relegion and Vegetarinism

Often our dietary choices are made up by relegion/relegious backgraound we were raised in.Some relegion prohibits eating certain/all meats or plants(Jainism-root based vegetables). Analyzing the relationship

between religion and vegetarianism highlights the ethical dimensions of dietary choices and how they are influenced by religious teachings.

```r
# Read the data
pl_relegion <- read_dta("pl.dta", col_select = c("pid", "syear", "ple0182", "plh0182","plh0258_h"))

# Exclusion values
exclusion_values <- c(-1, -2, -3, -4, -5, -6, -7, -8)

# Remove rows with exclusion values in plb0592 column
pl_relegion <- pl_relegion[!pl_relegion$ple0182 %in% exclusion_values, ]

# Remove rows with exclusion values in plb0585 column
pl_relegion <- pl_relegion[!pl_relegion$plh0258_h %in% exclusion_values, ]

# Map religion and diet preference codes to their respective labels
pl_relegion$ple0182 <- factor(pl_relegion$ple0182,
                              levels = c(1, 2, 3),
                              labels = c("Vegetarian", "Vegan", "Meat Eater"))

pl_relegion$plh0258_h <- factor(pl_relegion$plh0258_h,
                                levels = 1:11,
                                labels = c("Katholisch",
                                           "Evangelisch",
                                           "Andere christlische Religionsgemeinschaft",
                                           "Islamische Religionsgemeinschaft",
                                           "Andere Religionsgemeinschaft",
                                           "Konfessionslos",
                                           "Christlich orthodox",
                                           "Schiitische Rel.gemeinschaft",
                                           "Sunnitische Rel.gemeinschaft",
                                           "Alevitische Rel.gemeinschaft",
                                           "Mehrfachnennung"))

# Create a bar plot
ggplot(pl_relegion, aes(x = plh0258_h, fill = ple0182)) +
  geom_bar(position = "dodge") +
  labs(title = "Diet Preference by Religion",
       x = "Religion",
       y = "Count",


       fill = "Diet Preference") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



From the graph, it is evident that meat eaters (represented by the turquoise bars) significantly outnumber vegans (represented by the red bars) across all religious groups. The highest concentration of meat eaters is found within the "Sunnitische Religionsgemeinschaft" (Sunni Religious Community), followed by "Katholisch" (Catholic) and "Konfessionslos" (Non-religious/No affiliation). These groups show a strong preference for a meat-based diet, with very few individuals identifying as vegan.

Vegans are sparsely represented across the religions, with a small presence in the "Konfessionslos" category and a minimal appearance in "Andere Religionsgemeinschaft" (Other Religious Communities). This suggests that veganism is relatively uncommon within most religious groups represented in this dataset.

Surprisingly,There are no complete vegetarians in this dataset.

Overall, the graph indicates that religious affiliation may influence dietary preferences, with a dominant trend toward meat consumption across all religious groups. The limited representation of vegans suggests that either veganism is less prevalent among individuals within these religious affiliations, or that other factors, such as cultural or societal norms within these groups, may also play a role in shaping diet preferences.

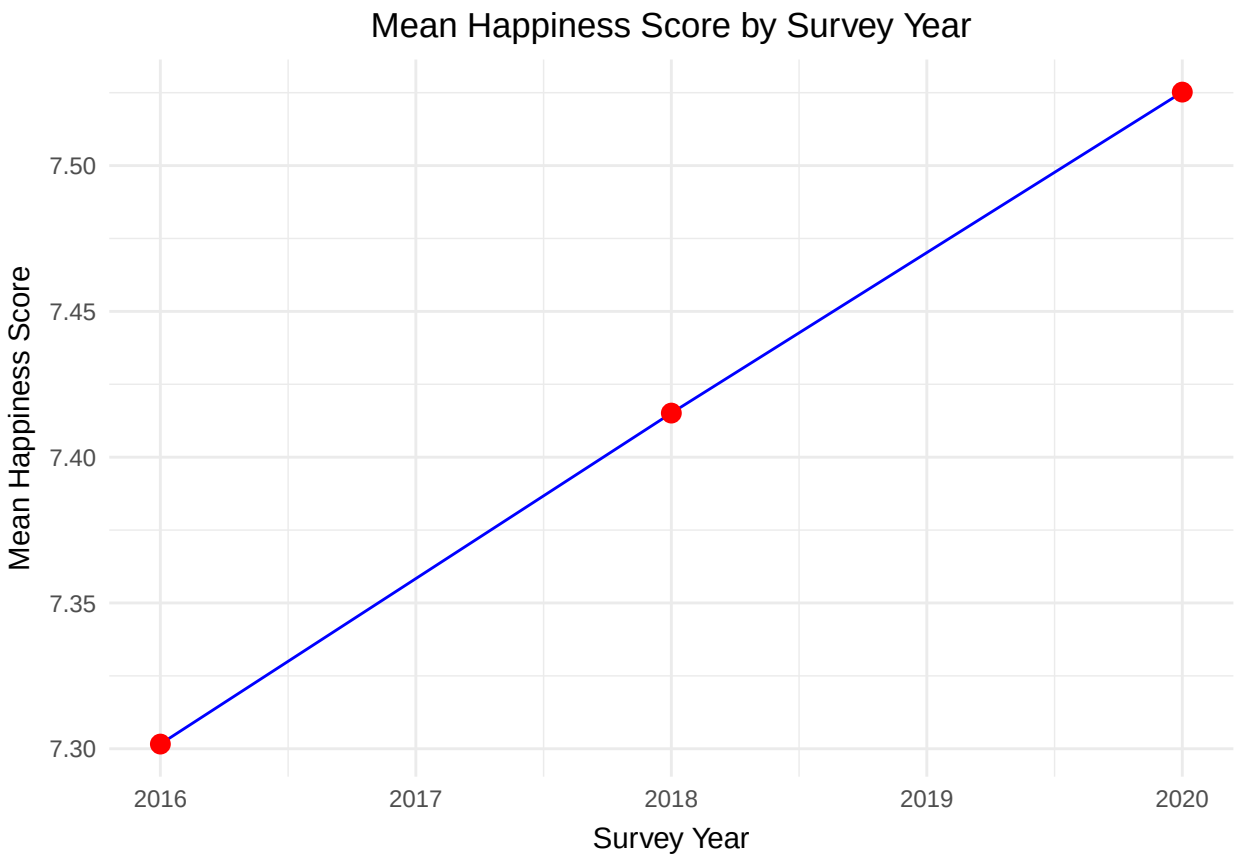
Outliers: The low representation of vegans and ***0 representation of vegetarians*** across most religions could be considered an outlier, but this is more a reflection of actual distribution rather than an anomalous data point.

3.2.3 Survey Year and Vegetarinism

This plot tries to find out if the mean happiness is increasing,while ***ceteris paribus(Other things remaining the same).***Analyzing how mean happiness changes over different survey years can help identify broader societal trends in well-being. If mean happiness is increasing or decreasing over time, it could reflect changes in societal conditions, economic factors, political stability, or cultural shifts that impact overall life satisfaction.

When analyzing mean happiness while holding other factors constant ***(ceteris paribus),*** you can isolate the effect of time itself on happiness. This helps determine whether happiness is inherently trending upwards or downwards over time, independent of other variables.

```
# Assuming merged_data already includes 'syear' and 'plh0182' (happiness scores)
# Group by survey year and calculate mean happiness
yearly_happiness <- merged_data %>%
  group_by(syear) %>%
  summarise(mean_happiness = mean(plh0182, na.rm = TRUE),  # Remove NA values for accurate calculation
            .groups = 'drop')  # This option drops the grouping after summarization


ggplot(yearly_happiness, aes(x = syear, y = mean_happiness)) +
  geom_line(group = 1, color = "blue") +  # Connect points with a line
  geom_point(color = "red", size = 3) +  # Add points to each year's mean happiness
  labs(title = "Mean Happiness Score by Survey Year",
       x = "Survey Year",
       y = "Mean Happiness Score") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5),
        axis.title.x = element_text(vjust = -0.2),
        axis.title.y = element_text(vjust = 2))
```



From the graph, it is evident that mean happiness scores have been steadily increasing over the period from 2016 to 2020. Starting from a mean happiness score of around 7.30 in 2016, there is a noticeable upward trend, reaching approximately 7.50 by 2020. This consistent rise suggests an overall improvement in

happiness levels among the surveyed population over these years.

The linear nature of the trend line indicates a fairly uniform increase in happiness, without significant fluctuations or deviations. This could imply that whatever factors are contributing to increased happiness, they have been steadily improving or consistently impacting the population in a positive manner during this period.

Outliers: There are no visible outliers, as the trend is linear and consistent without any abrupt changes or deviations.

3.2.4 Overall Satisfaction and Vegatrinism

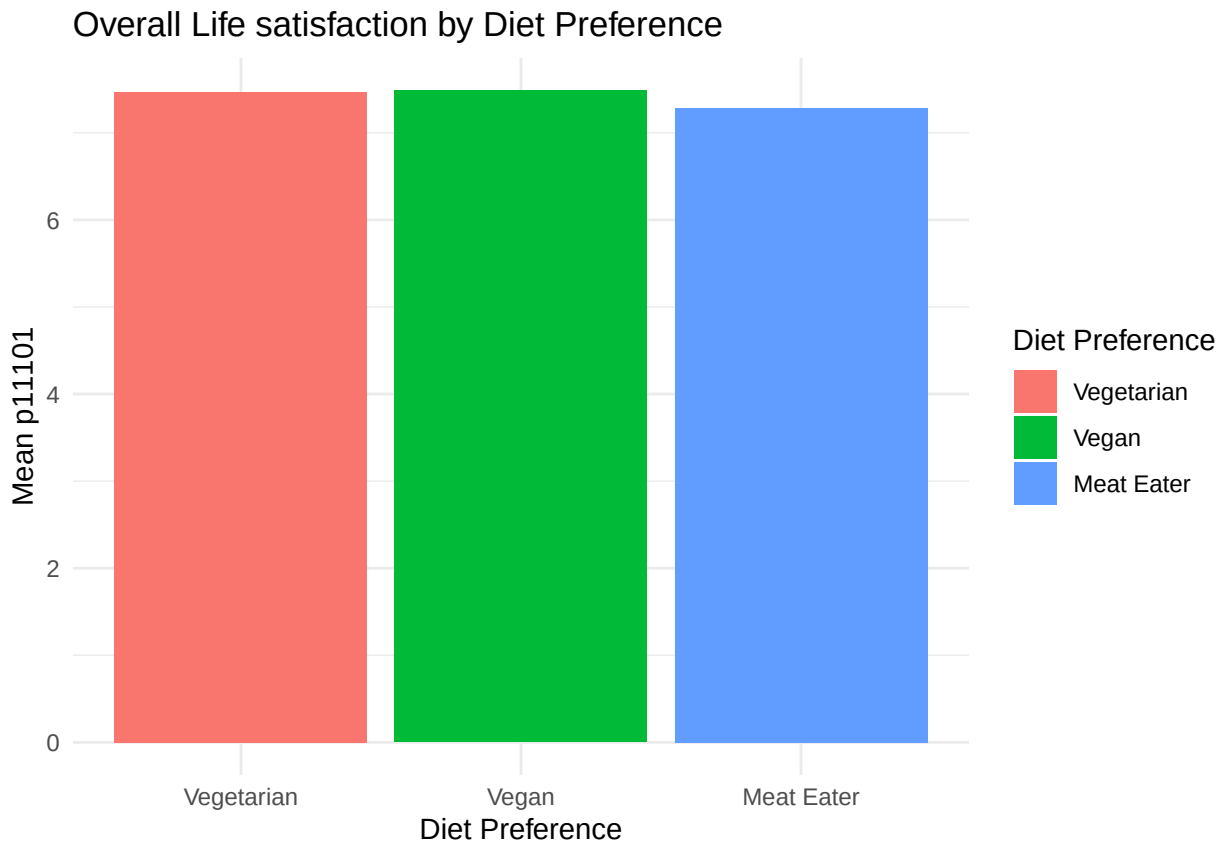
Inorder to see the validity of finidngs in Current life satisfaction,I took Overall life satisfaction for a sepertae analysis.If the relationship between vegetarianism and happiness remains consistent when using both "Current Life Satisfaction" and "Overall Life Satisfaction," this suggests that the impact of vegetarianism on happiness is stable over both short-term (current) and long-term (overall) perspectives. This consistency would strengthen the validity of findings, indicating that vegetarianism has a sustained impact on an individual's overall well-being.

```
# Load the data
pl_overall <- read_dta("pl.dta", col_select = c("pid", "syear", "ple0182", "plh0182"))
pequiv_overall <- read_dta("pequiv.dta", col_select = c("pid", "p11101"))

# Exclusion values
exclusion_values <- c(-1, -2, -3, -4, -5, -6, -7, -8)

pl_overall<- pl_overall[!pl_overall$ple0182 %in% exclusion_values, ]

# Remove rows with exclusion values in the m11126 column from pgen_income
pequiv_overall <- pequiv_overall[!pequiv_overall$p11101 %in% exclusion_values, ]

# Merge the education
merged_overall <- merge(pl_overall, pequiv_overall, by = "pid", all.x = TRUE)



merged_overall$ple0182 <- factor(merged_overall$ple0182,
                                 levels = c(1, 2, 3),
                                 labels = c("Vegetarian", "Vegan", "Meat Eater"))



# Calculate the mean p11101 for each diet preference
mean_values <- merged_overall %>%
  group_by(ple0182) %>%
  summarise(mean_p11101 = mean(p11101, na.rm = TRUE))

# Display the table
mean_values
```

```
## # A tibble: 3 x 2
##   ple0182    mean_p11101
##   <fct>            <dbl>
## 1 Vegetarian        7.47
```

```
## 2 Vegan             7.49
## 3 Meat Eater        7.29
```

```r
# Calculate the mean p11101 for each diet preference
mean_p11101 <- merged_overall %>%
  group_by(ple0182) %>%
  summarise(mean_p11101 = mean(p11101, na.rm = TRUE))

# Create a bar plot
ggplot(mean_p11101, aes(x = ple0182, y = mean_p11101, fill = ple0182)) +
  geom_bar(stat = "identity") +
  labs(title = "Overall Life satisfaction by Diet Preference",
       x = "Diet Preference",
       y = "Mean p11101",
       fill = "Diet Preference") +
  theme_minimal()
```



This bar chart illustrates the mean overall life satisfaction (labeled as p11101) across different diet preferences: Vegetarian (red), Vegan (green), and Meat Eater (blue). The x-axis categorizes the diet types, while the y-axis shows the mean overall life satisfaction score for each group.

From the graph, it is evident that vegetarians and vegans report slightly higher mean overall life satisfaction compared to meat eaters. The mean life satisfaction scores for vegetarians and vegans are nearly identical, both slightly higher than the score for meat eaters.

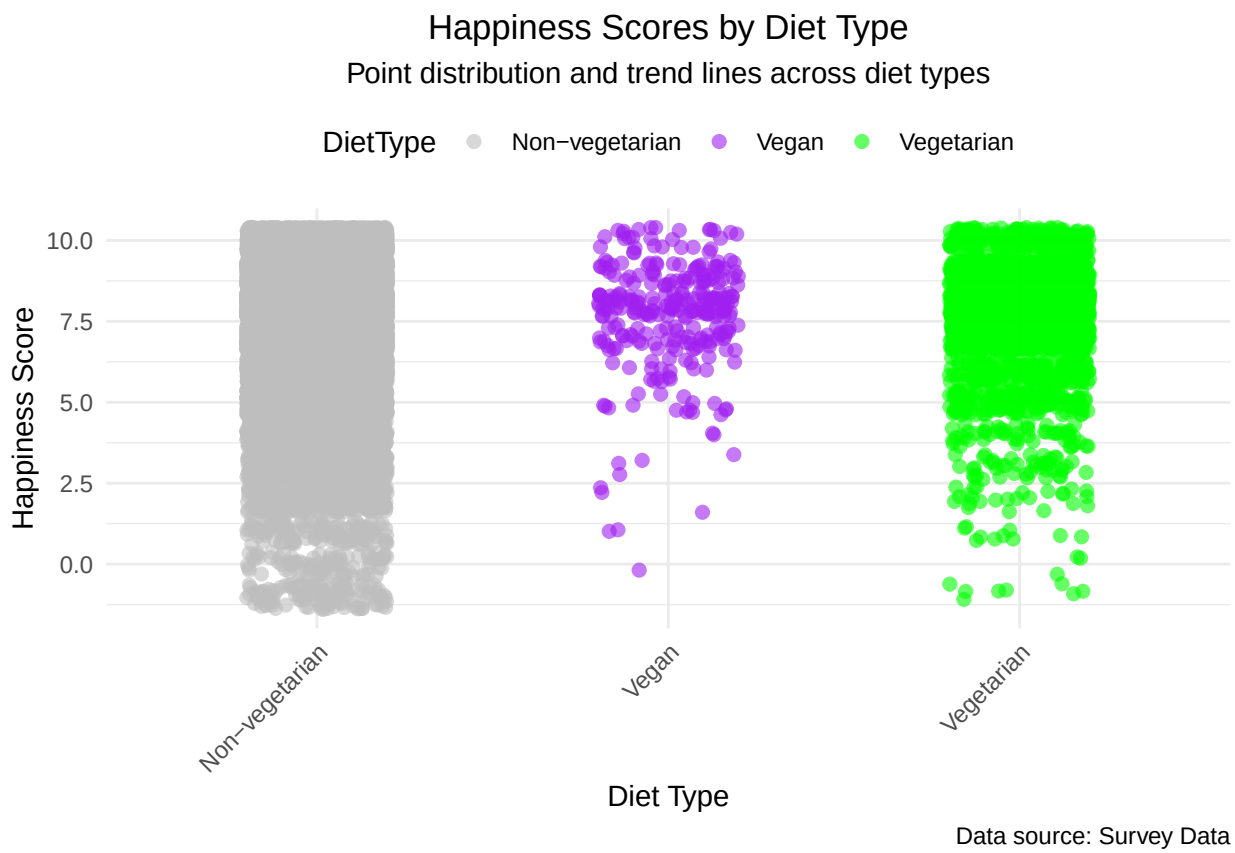
This suggests that, on average, individuals who follow a vegetarian or vegan diet tend to report higher overall life satisfaction than those who consume meat. The differences, however, are not drastic, indicating that

while diet preference might have some influence on life satisfaction, it may not be the sole determining factor. Other variables such as personal values, health, and lifestyle could also play significant roles in shaping an individual's overall sense of well-being.

Outliers: There are no visible outliers, as the scores are quite uniform across the diet groups.

3.3 Vegetarinism and Current-life satisfaction(Happiness)

```
# Assuming pl is loaded
# Transform pl to include DietType
pl <- pl %>%
  mutate(DietType = case_when(
    ple0182 == 1 ~ "Vegetarian",
    ple0182 == 2 ~ "Vegan",
    ple0182 == 3 ~ "Non-vegetarian"
  ))

# Filter for relevant diet types
pl_filtered <- pl[pl$ple0182 %in% c(1, 2, 3), ]

# Plotting with points and a smooth trend line
ggplot(pl_filtered, aes(x = DietType, y = pl_filtered$plh0182, color = DietType)) +
  geom_point(alpha = 0.6, position = position_jitter(width = 0.2), size = 2) +
  geom_smooth(method = "loess", se = FALSE, color = "black") +
  scale_color_manual(values = c("Vegetarian" = "green", "Vegan" = "purple", "Non-vegetarian" = "gray")) +
  labs(x = "Diet Type", y = "Happiness Score",
       title = "Happiness Scores by Diet Type",
       subtitle = "Point distribution and trend lines across diet types",
       caption = "Data source: Survey Data") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5),
        plot.subtitle = element_text(hjust = 0.5),
        plot.caption = element_text(hjust = 1),
        legend.position = "top",
        axis.text.x = element_text(angle = 45, hjust = 1))
```



This scatter plot shows the distribution of happiness scores across different diet types: Non-vegetarian (gray), Vegan (purple), and Vegetarian (green). The x-axis represents the diet types, while the y-axis indicates the happiness scores, which range from 0 to 10.

From the graph, it is evident that happiness scores vary widely within each diet group, but some patterns emerge:

Non-vegetarians (Gray): The happiness scores for non-vegetarians are spread across the entire range from 0 to 10, with a dense cluster between the mid to high range (around 5 to 10). This suggests that non-vegetarians experience a broad spectrum of happiness levels, though many report relatively high scores.

Vegans (Purple): Vegans also display a wide distribution of happiness scores, but with a noticeable concentration towards the higher end of the scale (around 7 to 10). This concentration indicates that a significant proportion of vegans report higher levels of happiness.

Vegetarians (Green): Similar to vegans, vegetarians tend to report higher happiness scores, with many scores clustering between 7 and 10. However, there is also a wider spread towards lower happiness scores compared to vegans, indicating more variability in how vegetarians perceive their happiness.

Overall, the graph suggests that while happiness scores are widely distributed within each diet group, vegans and vegetarians tend to report higher happiness levels more consistently than non-vegetarians. The clustering of higher scores among vegans and vegetarians might suggest that these diet types are associated with greater overall happiness, though individual experiences vary significantly within all groups.

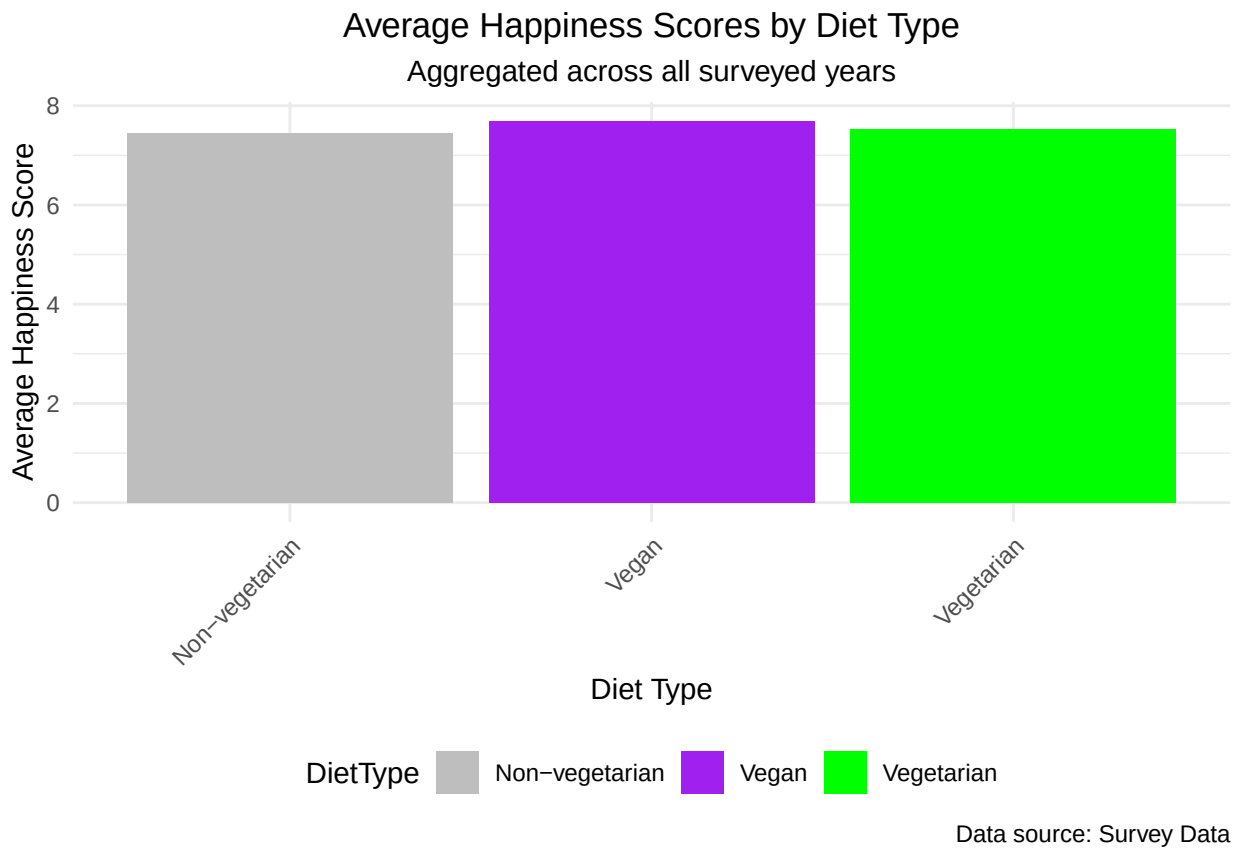
Outliers: Some very low happiness scores among non-vegetarians, especially given the wide spread, could be considered outliers, as well as some high happiness scores among individuals with low self-rated health.

```
# Group by DietType and calculate average happiness
average_happiness <- pl_filtered %>%
```

```r
  group_by(DietType) %>%
  summarise(AverageHappiness = mean(plh0182, na.rm = TRUE), .groups = 'drop')

# Print out the results
print(average_happiness)
```

```
## # A tibble: 3 x 2
##   DietType       AverageHappiness
##   <chr>                     <dbl>
## 1 Non-vegetarian             7.46
## 2 Vegan                      7.69
## 3 Vegetarian                 7.53
```

```r
# Plotting the average happiness levels
ggplot(average_happiness, aes(x = DietType, y = AverageHappiness, fill = DietType)) +
  geom_col() +
  scale_fill_manual(values = c("Vegetarian" = "green", "Vegan" = "purple", "Non-vegetarian" = "gray")) +
  labs(x = "Diet Type", y = "Average Happiness Score",
       title = "Average Happiness Scores by Diet Type",
       subtitle = "Aggregated across all surveyed years",
       caption = "Data source: Survey Data") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5),
        plot.subtitle = element_text(hjust = 0.5),
        plot.caption = element_text(hjust = 1),
        legend.position = "bottom",
        axis.text.x = element_text(angle = 45, hjust = 1))
```



From the graph, it is evident that all three diet types—Non-vegetarian, Vegan, and Vegetarian—have similar average happiness scores, with only minimal differences between them. Non-vegetarians have a slightly lower average happiness score compared to both vegans and vegetarians, who report nearly identical scores.

This suggests that, in general, diet type does not have a substantial impact on happiness levels. Whether an individual is a non-vegetarian, vegan, or vegetarian, their average happiness is relatively consistent. The similarity in scores implies that other factors beyond diet may be more influential in determining overall happiness, and that individuals following different dietary patterns experience comparable levels of well-being.

4 Overall Life satisfaction and Current Life satisfaction-assessment

1.Small Differences Across Diet Types:

The figures indicate only slight variations in happiness scores across the different diet types. This suggests that while there may be a positive association between vegetarianism/veganism and happiness, the effect size might be relatively modest. It implies that other factors beyond diet—such as social support, health, or lifestyle choices—might play more significant roles in influencing overall happiness.

2.Validity of Findings:

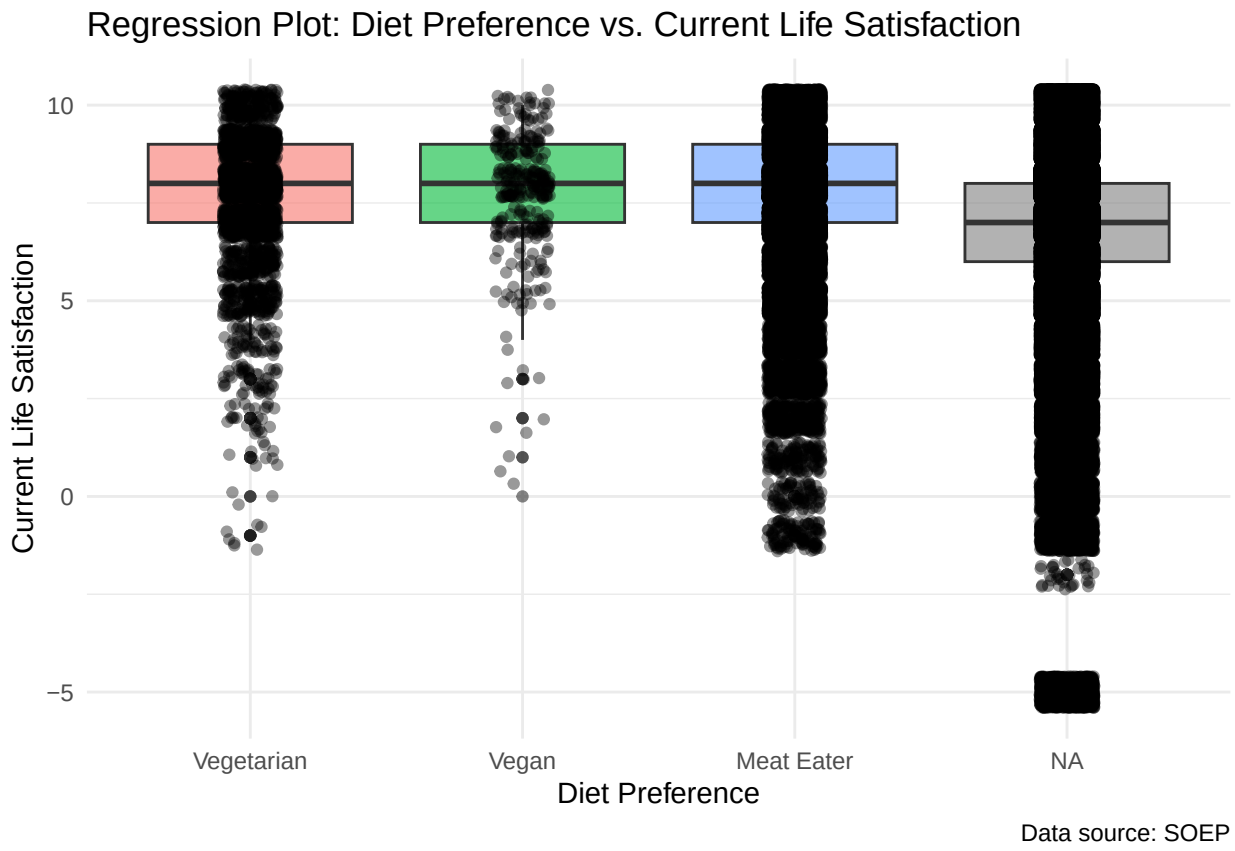
The consistency across both measures (current and overall life satisfaction) enhances the validity of your findings. It suggests that the observed relationship between diet and happiness is not an artifact of the time frame but rather a genuine, enduring effect.

3.Conclusion: The consistent findings between “Current Life Satisfaction” and “Overall Life Satisfaction” suggest that the positive impact of vegetarianism and veganism on happiness is stable over both short-term

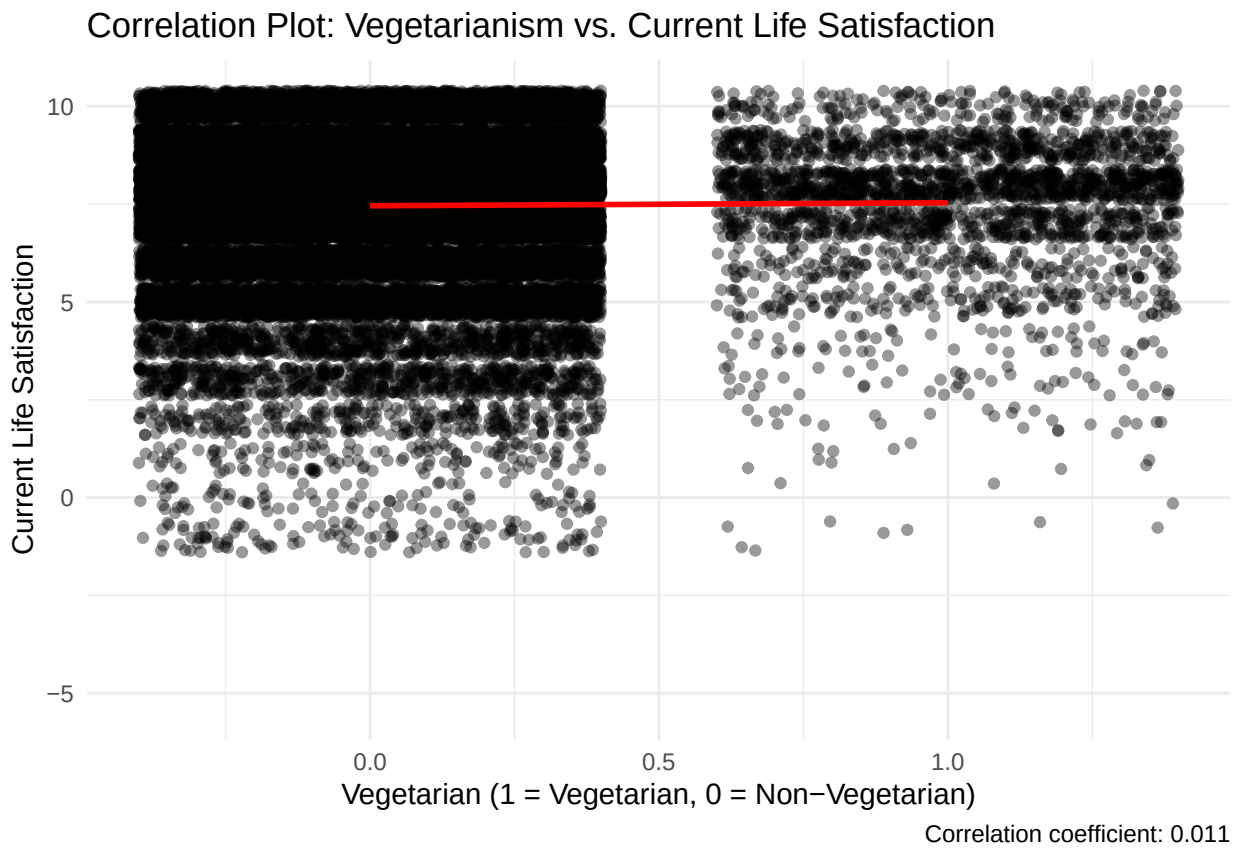
and long-term perspectives. The similarity in happiness scores across these two measures reinforces the validity of the relationship, indicating that these dietary choices contribute to sustained well-being. While the effect size appears modest, the consistency suggests that vegetarianism and veganism have a genuine, enduring positive influence on overall life satisfaction.

5 Regression and Co-relation.

The regression plot below illustrates the relationship between 'vegetarinism' and 'Happiness'. Each scatter point represents an individual observation from our estimation sample.




```
## [1] "Correlation coefficient: 0.0108671700092815"
```



A correlation coefficient of 0.011 suggests that there is a negligible positive linear relationship between 'vegetarianism' and 'current life satisfaction'. However, the strength of this relationship is extremely weak, implying that being vegetarian is only very slightly associated with differences in 'current life satisfaction'. This minimal correlation indicates that dietary preference alone does not significantly influence life satisfaction, and other factors likely play a more substantial role in determining overall happiness.

6 Presence of Outliers-Summary

Across the figures, there are a few potential outliers, particularly in the scatter plots where individuals report either significantly lower or higher happiness than expected based on their income or self-rated health. However, in the bar charts and trend lines, the data appears to be consistent with no significant outliers. The overall patterns suggest that while vegetarianism and veganism might have a positive influence on happiness, the effect size is modest, and other factors also play a significant role.

7 Limitations and Next Steps for In-Depth Analysis

7.1 Limitations

1.Despite the broader scope of this analysis, limitations remain, particularly in the depth of exploration into how these factors interact with each other. Future research should focus on a more detailed examination of the complex interplay between dietary choices and other life satisfaction determinants, potentially employing more advanced statistical techniques to capture these nuances.

2.Also,vegetarinism was taken into consideration in soep analysis only from 2016,and the data set where people responded in 1,2,3(Vegetarian,Vegan,Meat eating) preferences were low compariteveley.

3.Cross-Sectional Nature: The analysis is based on cross-sectional data, meaning it captures a snapshot in time rather than tracking changes over time. This limits the ability to draw causal inferences or understand how the relationships between variables might evolve.

4.Self-Reported Data: Variables like happiness and self-rated health are based on self-reported data, which can be subjective and prone to bias. Participants may overestimate or underestimate their happiness or health, leading to inaccuracies in the analysis. Simplistic Categorization:

5.The categorization of diet preferences (e.g., vegetarian, vegan, non-vegetarian) may oversimplify the complexity of individuals' dietary habits. For example, the "non-vegetarian" category includes a wide range of diets that might have different impacts on happiness and health. Potential Confounding Variables:

7.2 Next Steps for In-Depth Analysis:

1.Conduct Longitudinal Studies:

Implement a longitudinal study design to track changes in happiness, health, and dietary habits over time. This would provide insights into how sustained dietary choices impact long-term well-being. Include More Demographic and Psychosocial Variables:

2.Perform Multivariate Regression Analysis:

Use multivariate regression models to control for confounding variables and better isolate the effect of diet on happiness. This would help determine whether the observed associations hold when other factors are accounted for.

3.Examine Sub-Groups Within Diet Categories:

Break down the broad diet categories into more specific sub-groups (e.g., pescatarian, flexitarian, whole-food plant-based) to explore potential differences in their impact on happiness and health.

4.Qualitative Research:

Conduct qualitative research, such as interviews or focus groups, to explore participants' motivations for their dietary choices and how these relate to their perceptions of happiness and well-being. This could provide deeper insights into the psychological and social factors at play.

5.Investigate Causal Relationships:

Consider using experimental designs or instrumental variable approaches to establish causal relationships between diet and happiness, rather than merely observing associations.

6.Explore Cultural and Geographic Differences:

Expand the study to different cultural and geographic contexts to understand how cultural norms and regional dietary patterns influence the relationship between diet and happiness.

8 Conclusion

In conclusion, the analysis of the relationship between 'vegetarianism' and 'current life satisfaction' based on the SOEP dataset has provided valuable insights. Through data visualization, regression analysis, and correlation, I have identified an extremely weak positive correlation between 'vegetarianism' and 'current life satisfaction'. This suggests that, on average, individuals who follow a vegetarian diet may have a slightly different life satisfaction level compared to non-vegetarians. However, the correlation coefficient value of 0.011 indicates that this relationship is negligible and likely influenced by other factors not considered in this analysis.

In this project, I have also taken into consideration **8 additional variables**, including age, gender, income, education, health, religion, and overall life satisfaction, among others. The inclusion of these variables allowed for a more comprehensive analysis, revealing that while dietary preference may have some association with happiness, it is likely overshadowed by the influence of other socio-economic and demographic factors.

References

Aslanifar, E., Fakhri, M. K., Mirzaian, B., & Babaei Kafaki, H. (2014). *The comparison of personality traits and happiness of vegetarians and non-vegetarians*.

Beezhold, B. L., Johnston, C. S., & Daigle, D. R. (2010). Vegetarian diets are associated with healthy mood states: A cross-sectional study in seventh day adventist adults. *Nutrition Journal*, *9*(26). https://doi.org/10.1186/1475-2891-9-26

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., & Schupp, J. (2019). Board of directors SOEP and division head applied panel analysis in the german socio-economic panel study department. *Jahrbücher für Nationalökonomie Und Statistik*, *239*(2), 345–360. https://doi.org/10.1515/jbnst-2018-0022

Pfeiler, T. M., & Egloff, B. (2018). Examining the "veggie" personality: Results from a representative german sample. *Appetite*, *120*, 246–255. https://doi.org/10.1016/j.appet.2017.09.0052

Rehder, L. (2023). *Plant-based food goes mainstream in germany* (GM2023-0002). U.S. Department of Agriculture Foreign Agricultural Service. https://www.fas.usda.gov