```r
# Install R essentials
system("apt-get install --allow-releaseinfo-change -y r-base r-base-dev")
system("apt-get install --allow-releaseinfo-change -y libcurl4-openssl-dev libssl-dev libxml2-dev")

# Minimal setup with essential packages
if (!require(pacman)) install.packages("pacman")
pacman::p_load(tidyverse, haven, mice, car, nnet, lmtest)
```

```
Loading required package: pacman
```

```
Warning message in library(package, lib.loc = lib.loc, character.only = TRUE, logical.return = TRUE, :
“there is no package called ‘pacman’”
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

also installing the dependencies ‘rbibutils’, ‘Rdpack’, ‘minqa’, ‘nloptr’, ‘reformulas’, ‘ucminf’, ‘numDeriv’, ‘iterators’, ‘lme4’, ‘ordinal’, ‘foreach’, ‘shape’, ‘RcppEigen’, ‘pan’, ‘jomo’, ‘glmnet’, ‘mitml’



mice installed

Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

also installing the dependencies ‘cowplot’, ‘Deriv’, ‘microbenchmark’, ‘doBy’, ‘SparseM’, ‘MatrixModels’, ‘carData’, ‘abind’, ‘Formula’, ‘pbkrtest’, ‘quantreg’



car installed

Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

also installing the dependency ‘zoo’



lmtest installed
```

```r
# Install visualization-related packages
viz_packages <- c(
  'VIM',
  'visdat',
  'corrplot',
  'gridExtra'
)

# Install visualization packages
for(pkg in viz_packages) {
  if(!require(pkg, character.only = TRUE)) {
    install.packages(pkg)
    library(pkg, character.only = TRUE)
  }
}
```

```
Loading required package: VIM

Warning message in library(package, lib.loc = lib.loc, character.only
= TRUE, logical.return = TRUE, :
“there is no package called ‘VIM’”
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

also installing the dependencies ‘DEoptimR’, ‘proxy’, ‘robustbase’,
‘sp’, ‘vcd’, ‘e1071’, ‘laeken’, ‘ranger’


Loading required package: colorspace

Loading required package: grid

VIM is ready to use.


Suggestions and bug-reports can be submitted at:
https://github.com/statistikat/VIM/issues


Attaching package: ‘VIM’


The following object is masked from ‘package:datasets’:

    sleep


Loading required package: visdat

Warning message in library(package, lib.loc = lib.loc, character.only
= TRUE, logical.return = TRUE, :
```

```
“there is no package called ‘visdat’”
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

Loading required package: corrplot

Warning message in library(package, lib.loc = lib.loc, character.only
= TRUE, logical.return = TRUE, :
“there is no package called ‘corrplot’”
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

corrplot 0.95 loaded

Loading required package: gridExtra

Warning message in library(package, lib.loc = lib.loc, character.only
= TRUE, logical.return = TRUE, :
“there is no package called ‘gridExtra’”
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)


Attaching package: ‘gridExtra’


The following object is masked from ‘package:dplyr’:

    combine
```

```r
# Install statistical analysis packages
stat_packages <- c(
  'effectsize',
  'DescTools',
  'sandwich'
)

# Install statistical packages
for(pkg in stat_packages) {
  if(!require(pkg, character.only = TRUE)) {
    install.packages(pkg)
    library(pkg, character.only = TRUE)
  }
}
```

```
Loading required package: effectsize

Warning message in library(package, lib.loc = lib.loc, character.only
= TRUE, logical.return = TRUE, :
```

```
“there is no package called ‘effectsize’”
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

also installing the dependencies ‘bayestestR’, ‘insight’,
‘parameters’, ‘performance’, ‘datawizard’


Loading required package: DescTools

Warning message in library(package, lib.loc = lib.loc, character.only
= TRUE, logical.return = TRUE, :
“there is no package called ‘DescTools’”
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

also installing the dependencies ‘rootSolve’, ‘lmom’, ‘mvtnorm’,
‘expm’, ‘Exact’, ‘gld’



Attaching package: ‘DescTools’


The following object is masked from ‘package:car’:

    Recode


Loading required package: sandwich

Warning message in library(package, lib.loc = lib.loc, character.only
= TRUE, logical.return = TRUE, :
“there is no package called ‘sandwich’”
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)


# Load all required libraries
library(tidyverse)
library(haven)
library(ggplot2)
library(mice)
library(VIM)
library(forcats)
library(visdat)
library(car)
library(nnet)
library(lmtest)
library(sandwich)
library(gridExtra)
```

```r
library(vcd)
library(effectsize)
library(corrplot)
library(DescTools)

# Load the SPSS data file
# Update this path to match where you uploaded the file in your Google
Drive
qlfs_2015 <- read_sav("/content/lfsp_jm15_teaching_final2.sav")

# Remove CASENEW variable
qlfs_2015 <- qlfs_2015 %>%
  select(-CASENEW)

# Show initial data info
print("Dataset dimensions:")
print(dim(qlfs_2015))

print("Variable names:")
print(names(qlfs_2015))
```

```
[1] "Dataset dimensions:"
[1] 22428    13
[1] "Variable names:"
 [1] "PWT14"    "SEX"      "AGEEULR"  "MARSTA3R" "HIQUL15D" "ETHUK7R"
 [7] "ILODEFR"  "STAT3R"   "FTPTWK"   "TOTHRS"   "NSECMJ3R" "GOVTOF2"
[13] "PWT14R"
```

```r
# Install and load required packages for missing value analysis
if (!require(naniar)) install.packages("naniar")
if (!require(visdat)) install.packages("visdat")
if (!require(ggplot2)) install.packages("ggplot2")
library(naniar)
library(visdat)
library(ggplot2)

# Missing values analysis
missing_summary <- data.frame(
  Variable = names(qlfs_2015),
  Missing_Count = colSums(is.na(qlfs_2015)),
  Missing_Percentage = round(colSums(is.na(qlfs_2015)) /
nrow(qlfs_2015) * 100, 2)
)

# Display missing values summary
print("Missing Value Analysis:")
print(missing_summary)

# Create visualizations of missing values
# Plot 1: Missing values pattern using vis_miss
vis_miss(qlfs_2015) +
```

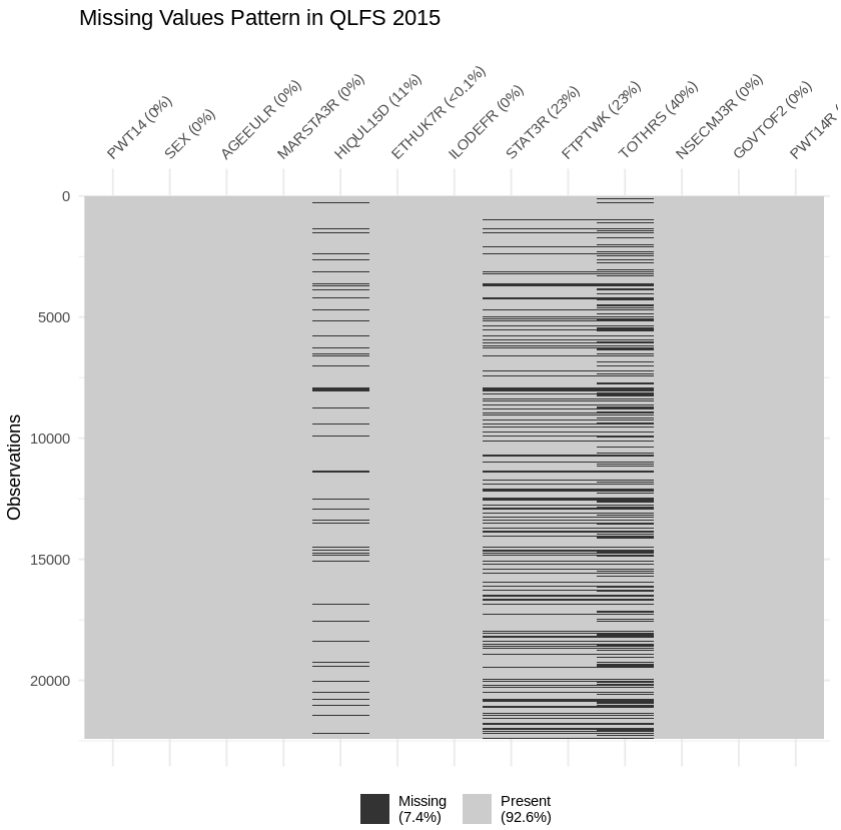
```
  ggtitle("Missing Values Pattern in QLFS 2015")

# Plot 2: Alternative missing values visualization using gg_miss_var
gg_miss_var(qlfs_2015) +
  theme_minimal() +
  labs(title = "Missing Values by Variable",
       x = "Variables",
       y = "Number of Missing Values")

# Plot 3: Missing values by variable (using basic ggplot)
ggplot(missing_summary, aes(x = reorder(Variable, Missing_Percentage),
y = Missing_Percentage)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  coord_flip() +
  theme_minimal() +
  labs(title = "Percentage of Missing Values by Variable",
       x = "Variables",
       y = "Missing Percentage (%)")
```

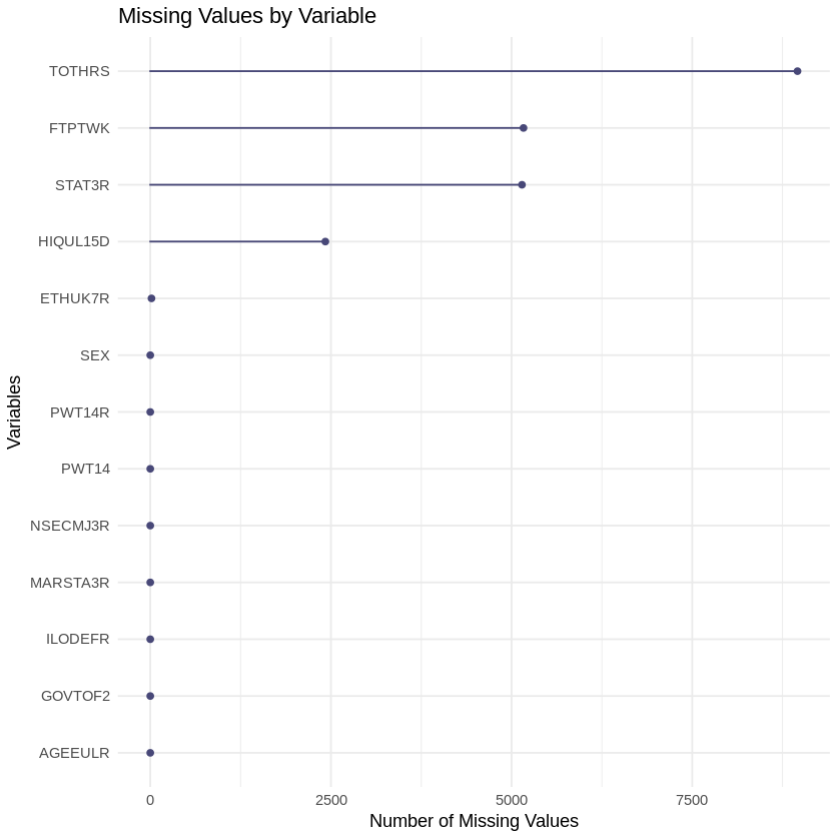
```
Loading required package: naniar
```

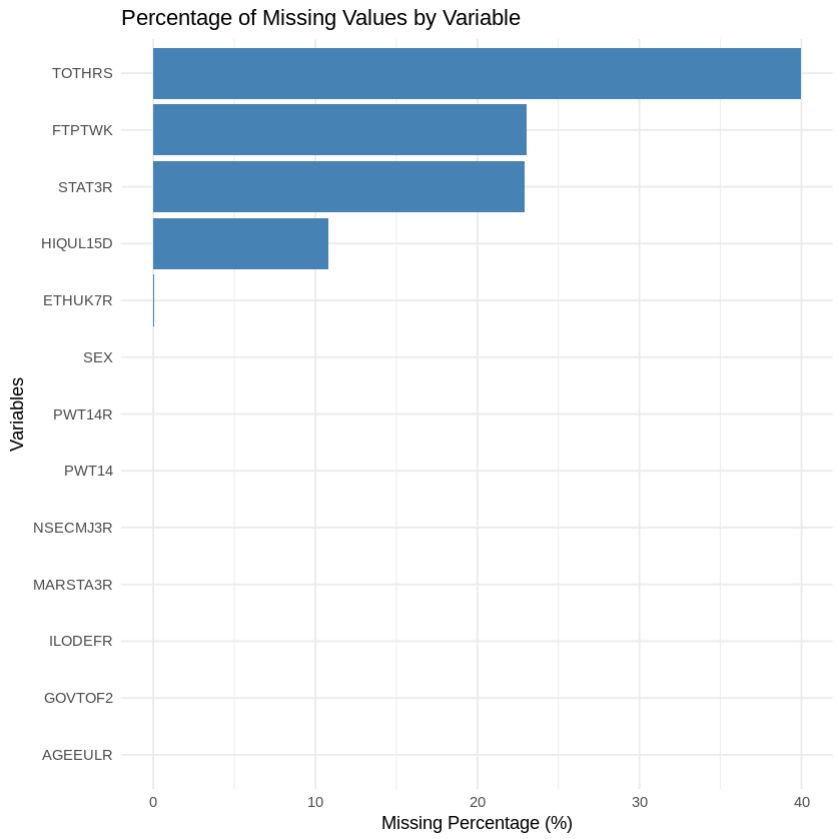
```
Warning message in library(package, lib.loc = lib.loc, character.only
= TRUE, logical.return = TRUE, :
“there is no package called ‘naniar’”
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)
```

```
also installing the dependencies ‘plyr’, ‘norm’, ‘viridis’, ‘UpSetR’
```

```
[1] "Missing Value Analysis:"
         Variable Missing_Count Missing_Percentage
PWT14       PWT14             0                0.00
SEX           SEX             0                0.00
AGEEULR   AGEEULR             0                0.00
MARSTA3R MARSTA3R             0                0.00
HIQUL15D HIQUL15D          2423               10.80
ETHUK7R   ETHUK7R            17                0.08
ILODEFR   ILODEFR             0                0.00
STAT3R     STAT3R          5144               22.94
FTPTWK     FTPTWK          5165               23.03
TOTHRS     TOTHRS          8957               39.94
NSECMJ3R NSECMJ3R             0                0.00
GOVTOF2   GOVTOF2             0                0.00
PWT14R     PWT14R             0                0.00
```


Missing Values Pattern in QLFS 2015
PWT14 (0%)
SEX (0%)
AGEEULR (0%)
MARSTA3R (0%)
HIQUL15D (11%)
ETHUK7R (<0.1%)
ILODEFR (0%)
STAT3R (23%)
FTPTWK (23%)
TOTHRS (40%)
NSECMJ3R (0%)
GOVTOF2 (0%)
0
5000
10000
15000
20000
Observations
Missing (7.4%)
Present (92.6%)


Missing Values by Variable
TOTHRS
FTPTWK
STAT3R
HIQUL15D
ETHUK7R
SEX
PWT14R
PWT14
NSECMJ3R
MARSTA3R
ILODEFR
GOVTOF2
AGEEULR
Variables
0
2500
5000
7500
Number of Missing Values




```r
# Define variable types
numeric_vars <- c("PWT14", "TOTHRS", "PWT14R")
categorical_vars <- c("SEX", "MARSTA3R", "HIQUL15D", "ETHUK7R",
"ILODEFR",
                      "STAT3R", "FTPTWK", "NSECMJ3R", "GOVTOF2")
ordinal_vars <- c("AGEEULR")

# Function to get mode for categorical variables
get_mode <- function(x) {
  if (all(is.na(x))) return(NA)
  ux <- unique(x[!is.na(x)])
  ux[which.max(tabulate(match(x[!is.na(x)], ux)))]
}

# Clean and transform variables
qlfs_2015_clean <- qlfs_2015 %>%
  # Replace special missing values
  mutate(across(everything(), ~if_else(. %in% c(-8, -9), NA_real_,
as.numeric(.)))) %>%
  # Convert variables to appropriate types with labels
  mutate(
    # Weights
    PWT14 = as.numeric(PWT14),
    PWT14R = as.numeric(PWT14R),
```

```
    # Demographic variables
    SEX = factor(SEX, levels = c(1, 2),
                 labels = c("Male", "Female")),

    MARSTA3R = factor(MARSTA3R, levels = 1:3,
                      labels = c("Single, never married",
                                 "Married/Cohabiting/Civil Partner",
                                 "Divorced/Widowed/Previously in Civil
Partnership")),

    AGEEULR = factor(AGEEULR, levels = 1:12,
                      labels = c("15-19", "20-24", "25-29", "30-34",
"35-39",
                                 "40-44", "45-49", "50-54", "55-59", "60-
64",
                                 "65-69", "70 and over")),

    # Education and ethnicity
    HIQUL15D = factor(HIQUL15D, levels = 1:7,
                       labels = c("Degree or equivalent", "Higher
education",
                                  "GCE, A-level or equivalent", "GCSE
grades A*-C or equivalent",
                                  "Other qualifications", "No
qualification", "Don't know")),

    ETHUK7R = factor(ETHUK7R, levels = 1:7,
                      labels = c("White", "Mixed/Multiple ethnic
groups", "Indian",
                                 "Pakistani or Bangladeshi",
                                 "Chinese or any other Asian background",
                                 "Black/African/Caribbean/Black British",
                                 "Other ethnic group")),

    # Employment variables
    ILODEFR = factor(ILODEFR, levels = 1:4,
                      labels = c("In employment", "ILO unemployed",
                                 "Inactive", "Under 16")),

    STAT3R = factor(STAT3R, levels = 1:3,
                    labels = c("Employee", "Self-employed",
                               "Government scheme or unpaid family
worker")),

    FTPTWK = factor(FTPTWK, levels = 1:2,
                    labels = c("Full-time", "Part-time")),

    TOTHRS = as.numeric(if_else(TOTHRS == 97, NA_real_,
as.numeric(TOTHRS))), # Handle 97+ hours special case
```

```
    NSECMJ3R = factor(NSECMJ3R, levels = 1:4,
                      labels = c("Higher managerial, administrative
and professional occupations",
                                 "Intermediate occupations and small
employers",
                                 "Routine and manual occupations",
                                 "Never worked, unemployed, and nec")),

    # Geographic variable
    GOVTOF2 = factor(GOVTOF2, levels = 1:13,
                     labels = c("North East", "North West (inc
Merseyside)",
                                "Yorkshire and Humberside", "East
Midlands",
                                "West Midlands", "Eastern", "London",
                                "South East", "South West", "Wales",
                                "Scotland", "Northern Ireland", "Does
not apply"))
  )

# Display sample of cleaned data
head(qlfs_2015_clean)
```

```
  PWT14 SEX    AGEEULR MARSTA3R

1 628   Male   45-49   Married/Cohabiting/Civil Partner

2 491   Female 45-49   Single, never married

3 641   Male   25-29   Single, never married

4 696   Female 55-59   Married/Cohabiting/Civil Partner

5 584   Female 65-69   Divorced/Widowed/Previously in Civil
Partnership
6 469   Female 40-44   Married/Cohabiting/Civil Partner

  HIQUL15D                          ETHUK7R ILODEFR       STAT3R   FTPTWK

1 GCSE grades A*-C or equivalent    White   In employment Employee Full-
time
2 Degree or equivalent              White   Inactive      Employee Full-
time
3 GCE, A-level or equivalent        White   In employment Employee Full-
time
4 No qualification                  White   Inactive      NA       NA

5 GCSE grades A*-C or equivalent    White   In employment Employee Full-
time
6 Degree or equivalent              White   Inactive      Employee Full-
```

```
time
  TOTHRS NSECMJ3R

1  0     Higher managerial, administrative and professional
occupations
2 NA     Higher managerial, administrative and professional
occupations
3 40     Routine and manual occupations

4 NA     Never worked, unemployed, and nec

5  0     Intermediate occupations and small employers

6 NA     Higher managerial, administrative and professional
occupations
  GOVTOF2       PWT14R
1 South West    0.9139198
2 South West    0.7145456
3 Wales         0.9328385
4 West Midlands 1.0128793
5 Scotland      0.8498872
6 South East    0.6825293
```

```r
# Perform imputation
qlfs_2015_imputed <- qlfs_2015_clean %>%
  mutate(
    # Numeric variables: median imputation
    across(all_of(numeric_vars), ~if_else(is.na(.), median(., na.rm =
TRUE), .)),

    # Categorical and ordinal variables: mode imputation
    across(all_of(c(categorical_vars, ordinal_vars)), function(x) {
      if_else(is.na(x),
              as.character(get_mode(x)),
              as.character(x)) %>%
        factor(levels = levels(x), labels = levels(x))
    })
  )

# Verify the imputation worked
print("Checking variable types after imputation:")
print(sapply(qlfs_2015_imputed, class))

# Check missing values after imputation
print("Missing values after imputation:")
print(colSums(is.na(qlfs_2015_imputed)))
```

```
[1] "Checking variable types after imputation:"
    PWT14       SEX   AGEEULR  MARSTA3R  HIQUL15D   ETHUK7R   ILODEFR
STAT3R
```

```
"numeric"  "factor"  "factor"  "factor"  "factor"  "factor"  "factor"
"factor"
   FTPTWK    TOTHRS  NSECMJ3R   GOVTOF2    PWT14R
 "factor" "numeric"  "factor"  "factor" "numeric"
[1] "Missing values after imputation:"
   PWT14      SEX  AGEEULR MARSTA3R HIQUL15D  ETHUK7R  ILODEFR
STAT3R
       0        0        0        0        0        0        0
0
  FTPTWK   TOTHRS NSECMJ3R  GOVTOF2   PWT14R
       0        0        0        0        0
```

```
if (!require(haven)) install.packages("haven")
library(tidyverse)
library(haven)

# Create frequency tables for categorical variables
cat_stats <- lapply(categorical_vars, function(var) { # Changed
cat_vars to categorical_vars
  freq_table <- table(qlfs_2015_imputed[[var]])
  prop_table <- prop.table(freq_table) * 100
  data.frame(
    Variable = var,
    Category = names(freq_table),
    Frequency = as.vector(freq_table),
    Percentage = round(as.vector(prop_table), 2)
  )
})

# Display categorical statistics
cat_stats_df <- do.call(rbind, cat_stats)
print(cat_stats_df)
```

```
   Variable
Category
1       SEX
Male
2       SEX
Female
3  MARSTA3R                                          Single, never
married
4  MARSTA3R                                Married/Cohabiting/Civil
Partner
5  MARSTA3R              Divorced/Widowed/Previously in Civil
Partnership
6  HIQUL15D                                           Degree or
equivalent
7  HIQUL15D                                              Higher
education
8  HIQUL15D                                     GCE, A-level or
```

```
equivalent
9  HIQUL15D                                    GCSE grades A*-C or
equivalent
10 HIQUL15D                                              Other
qualifications
11 HIQUL15D                                                 No
qualification
12 HIQUL15D                                                      Don't
know
13  ETHUK7R
White
14  ETHUK7R                                   Mixed/Multiple ethnic
groups
15  ETHUK7R
Indian
16  ETHUK7R                                      Pakistani or
Bangladeshi
17  ETHUK7R                           Chinese or any other Asian
background
18  ETHUK7R                           Black/African/Caribbean/Black
British
19  ETHUK7R                                        Other ethnic
group
20  ILODEFR                                          In
employment
21  ILODEFR                                        ILO
unemployed
22  ILODEFR
Inactive
23  ILODEFR
Under 16
24   STAT3R
Employee
25   STAT3R                                          Self-
employed
26   STAT3R                      Government scheme or unpaid family
worker
27   FTPTWK                                               Full-
time
28   FTPTWK                                               Part-
time
29 NSECMJ3R Higher managerial, administrative and professional
occupations
30 NSECMJ3R                    Intermediate occupations and small
employers
31 NSECMJ3R                                Routine and manual
occupations
32 NSECMJ3R                              Never worked, unemployed, and
nec
```

```
33  GOVTOF2                                                    North
East
34  GOVTOF2                                        North West (inc
Merseyside)
35  GOVTOF2                                           Yorkshire and
Humberside
36  GOVTOF2                                                East
Midlands
37  GOVTOF2                                                West
Midlands
38  GOVTOF2
Eastern
39  GOVTOF2
London
40  GOVTOF2                                                South
East
41  GOVTOF2                                                South
West
42  GOVTOF2
Wales
43  GOVTOF2
Scotland
44  GOVTOF2                                            Northern
Ireland
45  GOVTOF2                                              Does not
apply
   Frequency Percentage
1      10656      47.51
2      11772      52.49
3       7179      32.01
4      12303      54.86
5       2946      13.14
6       7647      34.10
7       1937       8.64
8       4539      20.24
9       4195      18.70
10      1820       8.11
11      2096       9.35
12       194       0.86
13     20168      89.92
14       201       0.90
15       495       2.21
16       436       1.94
17       294       1.31
18       515       2.30
19       319       1.42
20     13751      61.31
21       754       3.36
22      7923      35.33
```

```
23          0       0.00
24      19962      89.00
25       2385      10.63
26         81       0.36
27      17164      76.53
28       5264      23.47
29       6922      30.86
30       4029      17.96
31       5448      24.29
32       6029      26.88
33        976       4.35
34       2478      11.05
35          0       0.00
36       1982       8.84
37       1624       7.24
38       1870       8.34
39       2080       9.27
40       2387      10.64
41       3113      13.88
42       1970       8.78
43       1125       5.02
44       1914       8.53
45        909       4.05
```

```r
# Calculate summary statistics for numerical variables
num_stats <- lapply(numeric_vars, function(var) {
  data.frame(
    Variable = var,
    Mean = mean(qlfs_2015_imputed[[var]], na.rm = TRUE),
    Median = median(qlfs_2015_imputed[[var]], na.rm = TRUE),
    SD = sd(qlfs_2015_imputed[[var]], na.rm = TRUE),
    Min = min(qlfs_2015_imputed[[var]], na.rm = TRUE),
    Max = max(qlfs_2015_imputed[[var]], na.rm = TRUE)
  )
})

# Display numerical statistics
num_stats_df <- do.call(rbind, num_stats)
print(num_stats_df)
```

```
  Variable       Mean      Median          SD Min        Max
1    PWT14 687.152666 622.0000000 403.0655576   0 7442.00000
2   TOTHRS  33.750178  37.0000000  13.7592561   0   96.00000
3   PWT14R   1.000004   0.9051881   0.5865758   0   10.83024
```

```r
# Create cross-tabulations for key variables
key_cross_tabs <- list(
  "Employment Status by Region" = table(qlfs_2015_imputed$STAT3R,
qlfs_2015_imputed$GOVTOF2),
  "Employment Status by Education" = table(qlfs_2015_imputed$STAT3R,
```

```
qlfs_2015_imputed$HIQUL15D),
  "Employment Status by Socio-economic Class" =
table(qlfs_2015_imputed$STAT3R, qlfs_2015_imputed$NSECMJ3R)
)

# Display cross-tabulations
for (name in names(key_cross_tabs)) {
  cat("\n", name, "\n")
  print(key_cross_tabs[[name]])
  cat("\n Chi-square test:\n")
  print(chisq.test(key_cross_tabs[[name]]))
  cat("\n")
}


 Employment Status by Region

                                                North East
  Employee                                             904
  Self-employed                                         70
  Government scheme or unpaid family worker              2

                                                North West (inc
Merseyside)
  Employee
2230
  Self-employed
243
  Government scheme or unpaid family worker
5

                                                Yorkshire and Humberside
  Employee                                                             0
  Self-employed                                                        0
  Government scheme or unpaid family worker                            0

                                                East Midlands West
Midlands Eastern
  Employee                                                  1776
1455    1690
  Self-employed                                              202
162     171
  Government scheme or unpaid family worker                    4
7       9

                                                London South East South
West Wales
  Employee                                        1840       2067
2719  1707
  Self-employed                                    232        313
```

```
386   257
  Government scheme or unpaid family worker      8          7
8     6

                                                Scotland Northern Ireland
  Employee                                          1011             1740
  Self-employed                                      106              165
  Government scheme or unpaid family worker            8                9

                                                Does not apply
  Employee                                                 823
  Self-employed                                             78
  Government scheme or unpaid family worker                  8

 Chi-square test:

Warning message in chisq.test(key_cross_tabs[[name]]):
“Chi-squared approximation may be incorrect”


     Pearson's Chi-squared test

data:  key_cross_tabs[[name]]
X-squared = NaN, df = 24, p-value = NA



 Employment Status by Education

                                                Degree or equivalent
  Employee                                                      6836
  Self-employed                                                  781
  Government scheme or unpaid family worker                       30

                                                Higher education
  Employee                                                  1660
  Self-employed                                              264
  Government scheme or unpaid family worker                   13

                                                GCE, A-level or equivalent
  Employee                                                           3987
  Self-employed                                                       538
  Government scheme or unpaid family worker                            14

                                                GCSE grades A*-C or
equivalent
  Employee
3793
  Self-employed
388
```

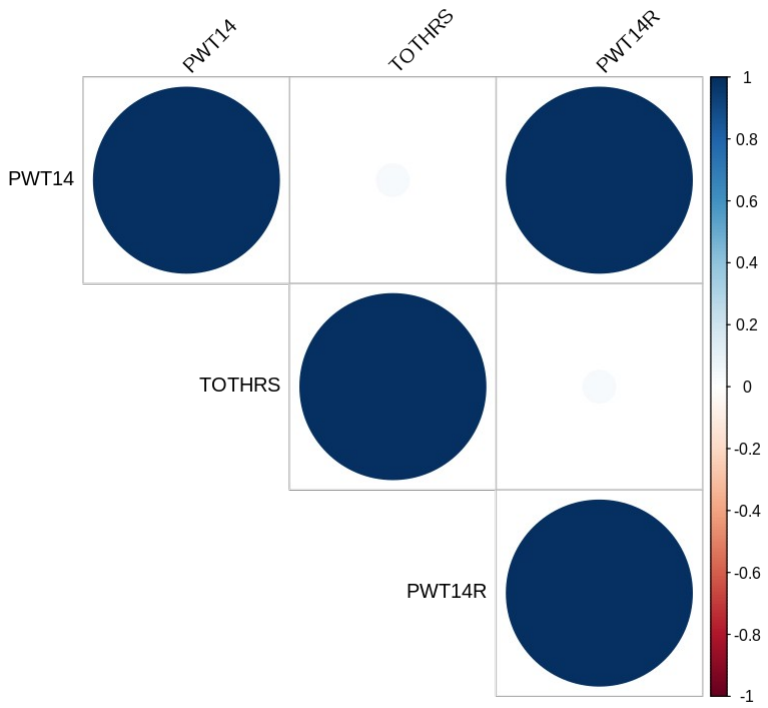
```
  Government scheme or unpaid family worker
14

                                                 Other qualifications
  Employee                                                       1620
  Self-employed                                                   198
  Government scheme or unpaid family worker                         2

                                                 No qualification Don't
know
  Employee                                                     1889
177
  Self-employed                                                 200
16
  Government scheme or unpaid family worker                       7
1

 Chi-square test:

Warning message in chisq.test(key_cross_tabs[[name]]):
“Chi-squared approximation may be incorrect”


     Pearson's Chi-squared test

data:  key_cross_tabs[[name]]
X-squared = 48.709, df = 12, p-value = 2.352e-06



 Employment Status by Socio-economic Class

                                                 Higher managerial,
administrative and professional occupations
  Employee
6279
  Self-employed
639
  Government scheme or unpaid family worker
4

                                                 Intermediate occupations
and small employers
  Employee
2319
  Self-employed
1707
  Government scheme or unpaid family worker
3
```

```
                                                  Routine and manual
occupations
  Employee
5439
  Self-employed
0
  Government scheme or unpaid family worker
9

                                                  Never worked, unemployed,
and nec
  Employee
5925
  Self-employed
39
  Government scheme or unpaid family worker
65

 Chi-square test:

     Pearson's Chi-squared test

data:  key_cross_tabs[[name]]
X-squared = 5671.4, df = 6, p-value < 2.2e-16
```

```r
# Calculate mean hours worked by various categories
hours_by_region <- aggregate(TOTHRS ~ GOVTOF2, data =
qlfs_2015_imputed, FUN = mean)
hours_by_status <- aggregate(TOTHRS ~ STAT3R, data =
qlfs_2015_imputed, FUN = mean)
hours_by_nsec <- aggregate(TOTHRS ~ NSECMJ3R, data =
qlfs_2015_imputed, FUN = mean)

# Display mean hours analysis
print("Mean hours worked by region:")
print(hours_by_region)

print("Mean hours worked by employment status:")
print(hours_by_status)

print("Mean hours worked by socio-economic class:")
print(hours_by_nsec)
```

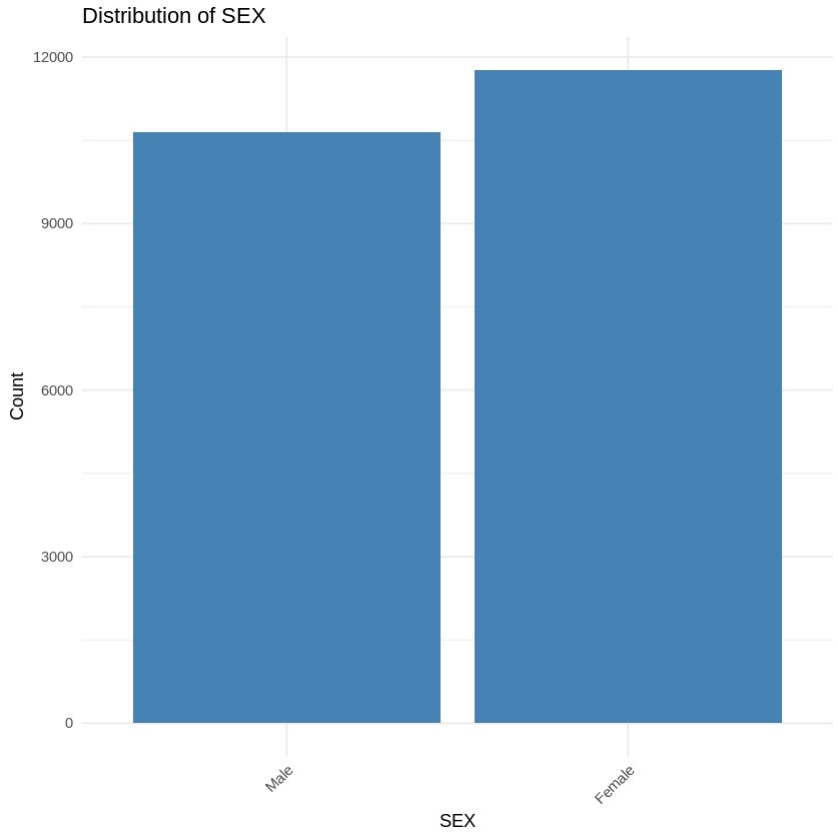
```
[1] "Mean hours worked by region:"
                       GOVTOF2   TOTHRS
1                   North East 32.92930
2  North West (inc Merseyside) 33.90759
3                East Midlands 33.37740
```

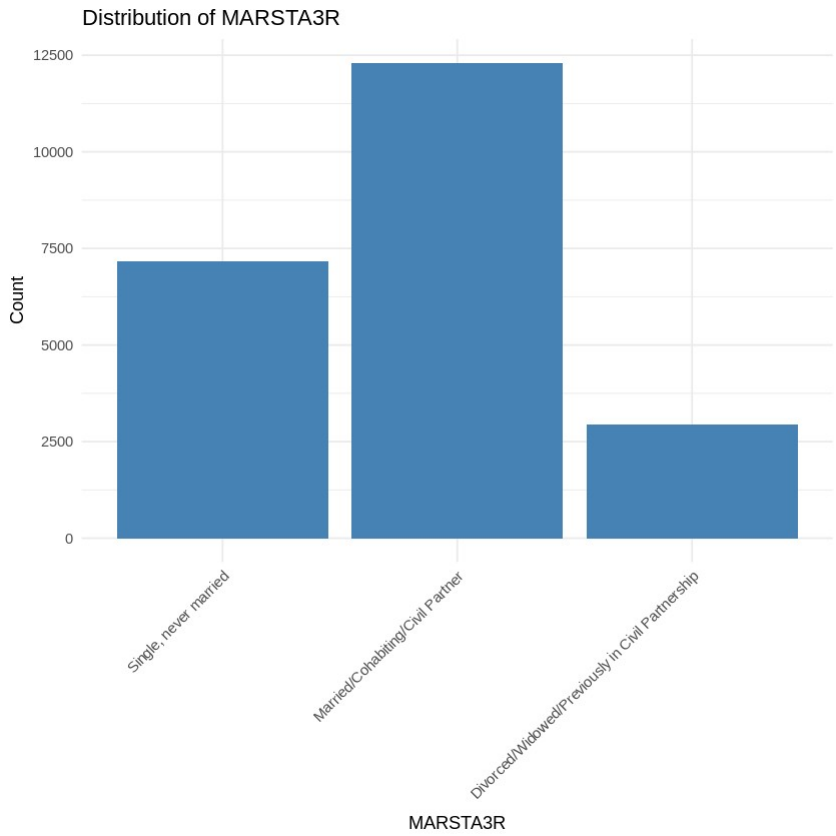
```
4                 West Midlands 34.52648
5                       Eastern 34.13316
6                        London 33.58365
7                    South East 34.22162
8                    South West 33.27369
9                         Wales 33.18426
10                     Scotland 33.93867
11            Northern Ireland 33.52299
12              Does not apply 35.08691
[1] "Mean hours worked by employment status:"
                                       STAT3R   TOTHRS
1                                    Employee 34.00371
2                                Self-employed 32.20545
3 Government scheme or unpaid family worker 16.75309
[1] "Mean hours worked by socio-economic class:"
                                                                NSECMJ3R
TOTHRS
1 Higher managerial, administrative and professional occupations
34.96331
2                     Intermediate occupations and small employers
31.33656
3                                    Routine and manual occupations
32.54112
4                                  Never worked, unemployed, and nec
35.06286

# Create correlation matrix for numeric variables
cor_matrix <- cor(qlfs_2015_imputed[, numeric_vars], use =
"complete.obs")

# Display correlation matrix
print("Correlation matrix:")
print(cor_matrix)

# Visualize correlation matrix
corrplot(cor_matrix, method = "circle", type = "upper",
         tl.col = "black", tl.srt = 45)

[1] "Correlation matrix:"
            PWT14     TOTHRS     PWT14R
PWT14  1.00000000 0.03076817 1.00000000
TOTHRS 0.03076817 1.00000000 0.03076817
PWT14R 1.00000000 0.03076817 1.00000000
```



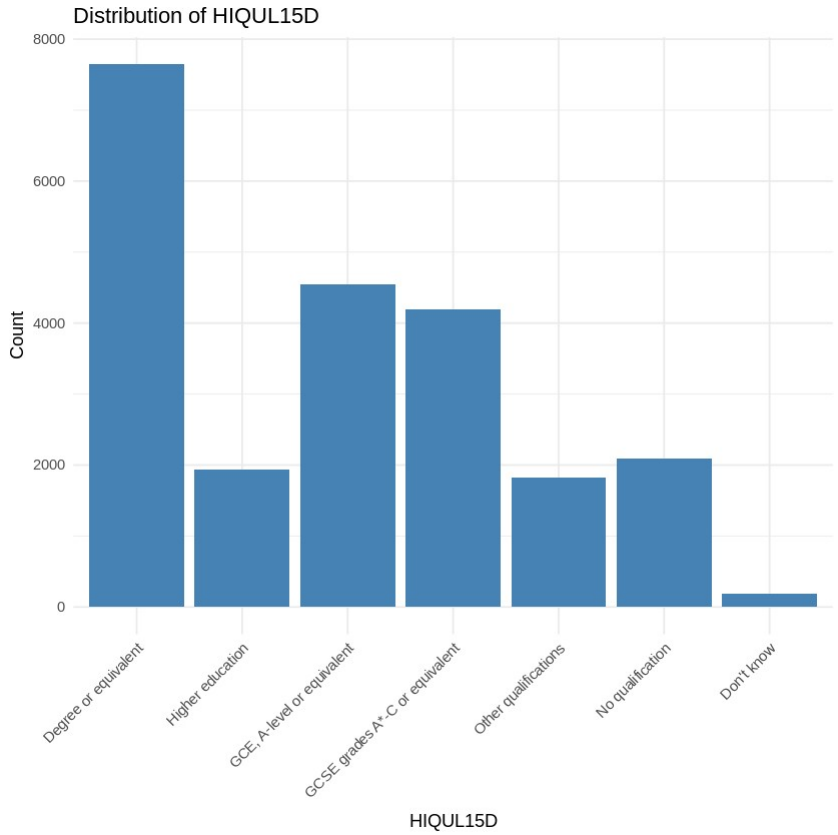

```r
# Function for categorical variable plots
plot_categorical <- function(data, var_name) {
  ggplot(data, aes(x = .data[[var_name]])) +
    geom_bar(fill = "steelblue") +
    theme_minimal() +
    theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
    labs(title = paste("Distribution of", var_name),
         x = var_name,
         y = "Count")
}

# Function for numerical variable plots
plot_numerical <- function(data, var_name) {
  ggplot(data, aes(x = .data[[var_name]])) +
    geom_histogram(fill = "steelblue", bins = 30) +
    geom_density(color = "red") +
    theme_minimal() +
    labs(title = paste("Distribution of", var_name),
         x = var_name,
         y = "Count")
}

# Create plots for categorical variables (display first 3 for brevity)
# Changed cat_vars to categorical_vars
```

```r
for (var in categorical_vars[1:3]) {
  print(plot_categorical(qlfs_2015_imputed, var))
}

# Create plots for numerical variables
for (var in num_vars) {
  print(plot_numerical(qlfs_2015_imputed, var))
}
```



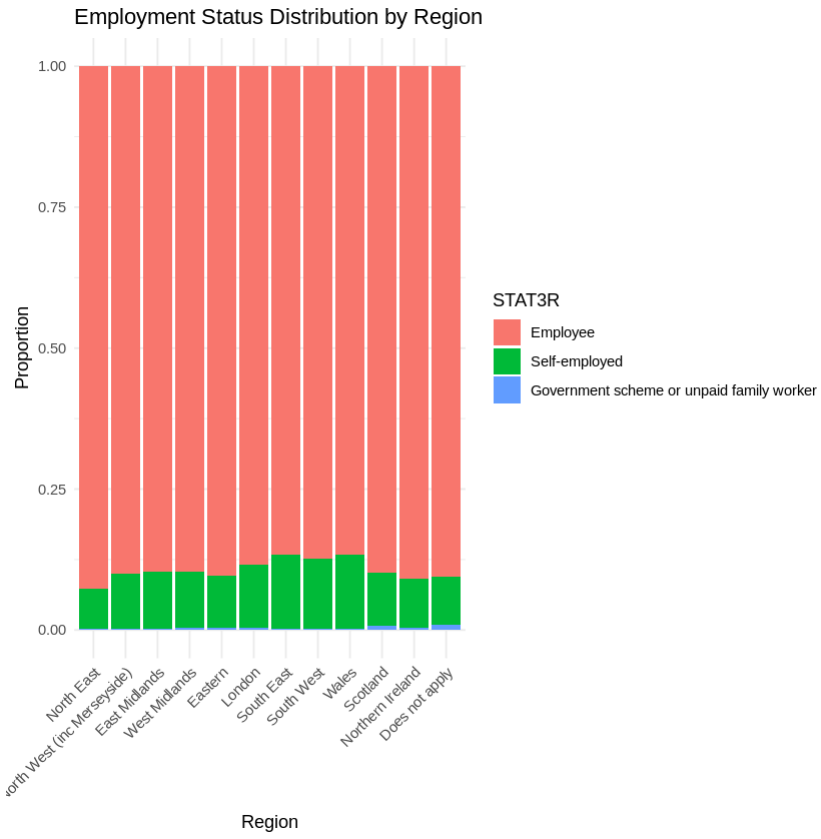


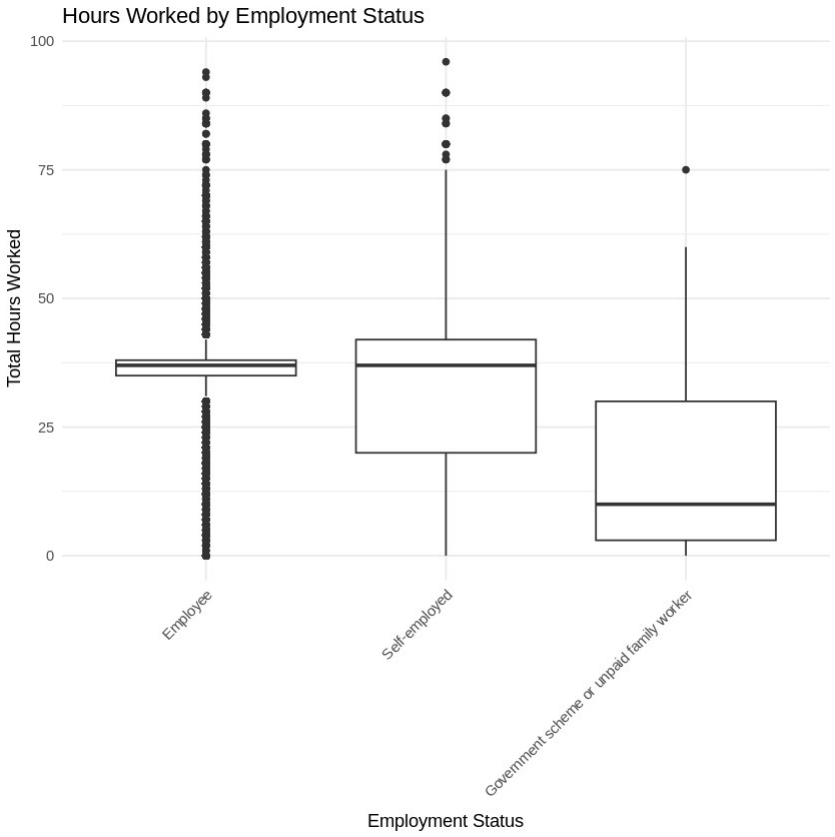

```
Error: object 'num_vars' not found
Traceback:
```

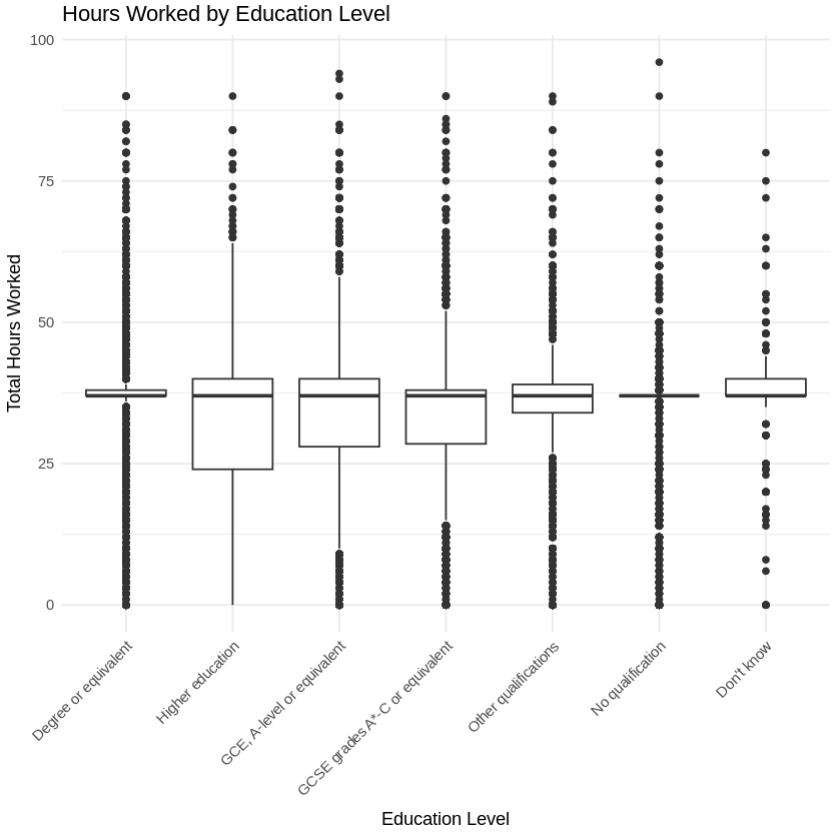



```r
# Employment Status by Region
ggplot(qlfs_2015_imputed, aes(x = GOVTOF2, fill = STAT3R)) +
  geom_bar(position = "fill") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title = "Employment Status Distribution by Region",
       x = "Region",
       y = "Proportion")

# Hours Worked by Employment Status
ggplot(qlfs_2015_imputed, aes(x = STAT3R, y = TOTHRS)) +
  geom_boxplot() +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title = "Hours Worked by Employment Status",
       x = "Employment Status",
       y = "Total Hours Worked")

# Hours Worked by Education Level
ggplot(qlfs_2015_imputed, aes(x = HIQUL15D, y = TOTHRS)) +
  geom_boxplot() +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title = "Hours Worked by Education Level",
```

```
       x = "Education Level",
       y = "Total Hours Worked")
```




Hours Worked by Employment Status
Total Hours Worked
100
75
50
25
0
Employee
Self-employed
Government scheme or unpaid family worker
Employment Status


Hours Worked by Education Level
Total Hours Worked
100
75
50
25
0
Degree or equivalent
Higher education
GCE, A-level or equivalent
GCSE grades A*-C or equivalent
Other qualifications
No qualification
Don't know
Education Level

```r
# Chi-square tests for key relationships
key_relationships <- list(
  "Employment Status vs Region" =
chisq.test(table(qlfs_2015_imputed$STAT3R,
qlfs_2015_imputed$GOVTOF2)),
  "Employment Status vs Education" =
chisq.test(table(qlfs_2015_imputed$STAT3R,
qlfs_2015_imputed$HIQUL15D)),
  "Employment Status vs Socio-economic Class" =
chisq.test(table(qlfs_2015_imputed$STAT3R,
qlfs_2015_imputed$NSECMJ3R))
)

# Display chi-square test results
for (name in names(key_relationships)) {
  cat("\n", name, "\n")
  print(key_relationships[[name]])
  cat("\n")
}
```

```
Warning message in chisq.test(table(qlfs_2015_imputed$STAT3R,
qlfs_2015_imputed$GOVTOF2)):
“Chi-squared approximation may be incorrect”
Warning message in chisq.test(table(qlfs_2015_imputed$STAT3R,
qlfs_2015_imputed$HIQUL15D)):
“Chi-squared approximation may be incorrect”


 Employment Status vs Region

     Pearson's Chi-squared test

data:  table(qlfs_2015_imputed$STAT3R, qlfs_2015_imputed$GOVTOF2)
X-squared = NaN, df = 24, p-value = NA



 Employment Status vs Education

     Pearson's Chi-squared test

data:  table(qlfs_2015_imputed$STAT3R, qlfs_2015_imputed$HIQUL15D)
X-squared = 48.709, df = 12, p-value = 2.352e-06



 Employment Status vs Socio-economic Class

     Pearson's Chi-squared test

data:  table(qlfs_2015_imputed$STAT3R, qlfs_2015_imputed$NSECMJ3R)
```

```
X-squared = 5671.4, df = 6, p-value < 2.2e-16
```

```
# ANOVA for hours worked by various categories
anova_tests <- list(
  "Hours by Employment Status" = aov(TOTHRS ~ STAT3R, data =
qlfs_2015_imputed),
  "Hours by Education" = aov(TOTHRS ~ HIQUL15D, data =
qlfs_2015_imputed),
  "Hours by Region" = aov(TOTHRS ~ GOVTOF2, data = qlfs_2015_imputed)
)

# Display ANOVA test results
for (name in names(anova_tests)) {
  cat("\n", name, "\n")
  print(summary(anova_tests[[name]]))
  cat("\n")
}
```

```
 Hours by Employment Status
              Df  Sum Sq Mean Sq F value Pr(>F)
STAT3R         2   30375   15188   80.79 <2e-16 ***
Residuals   22425 4215440     188
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
 Hours by Education
              Df  Sum Sq Mean Sq F value Pr(>F)
HIQUL15D       6   19867    3311   17.57 <2e-16 ***
Residuals   22421 4225948     188
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
 Hours by Region
              Df  Sum Sq Mean Sq F value   Pr(>F)
GOVTOF2       11    5936   539.7   2.853 0.000962 ***
Residuals   22416 4239879   189.1
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
# Tukey's HSD Tests
for (name in names(anova_tests)) {
  cat("\n", name, "\n")
  print(TukeyHSD(anova_tests[[name]]))
  cat("\n")
}
```

```
 Hours by Employment Status
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = TOTHRS ~ STAT3R, data = qlfs_2015_imputed)

$STAT3R
                                                                  diff
lwr
Self-employed-Employee                                    -1.798256  -
2.494481
Government scheme or unpaid family worker-Employee      -17.250621 -
20.828476
Government scheme or unpaid family worker-Self-employed -15.452364 -
19.083110
                                                                   upr p
adj
Self-employed-Employee                                    -1.102032
0
Government scheme or unpaid family worker-Employee      -13.672765
0
Government scheme or unpaid family worker-Self-employed -11.821619
0



 Hours by Education
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = TOTHRS ~ HIQUL15D, data = qlfs_2015_imputed)

$HIQUL15D
                                                                 diff
Higher education-Degree or equivalent                     -2.50228051
GCE, A-level or equivalent-Degree or equivalent           -1.64978866
GCSE grades A*-C or equivalent-Degree or equivalent       -1.97426495
Other qualifications-Degree or equivalent                 -0.73495991
No qualification-Degree or equivalent                     -0.82792821
Don't know-Degree or equivalent                            2.04109084
GCE, A-level or equivalent-Higher education                0.85249185
GCSE grades A*-C or equivalent-Higher education            0.52801557
Other qualifications-Higher education                      1.76732060
No qualification-Higher education                          1.67435231
Don't know-Higher education                                4.54337135
GCSE grades A*-C or equivalent-GCE, A-level or equivalent -0.32447628
Other qualifications-GCE, A-level or equivalent            0.91482875
No qualification-GCE, A-level or equivalent                0.82186046
Don't know-GCE, A-level or equivalent                      3.69087950
```

```
Other qualifications-GCSE grades A*-C or equivalent        1.23930503
No qualification-GCSE grades A*-C or equivalent            1.14633674
Don't know-GCSE grades A*-C or equivalent                  4.01535579
No qualification-Other qualifications                     -0.09296829
Don't know-Other qualifications                            2.77605075
Don't know-No qualification                                2.86901904
                                                                  lwr
Higher education-Degree or equivalent                     -3.53198261
GCE, A-level or equivalent-Degree or equivalent           -2.40828502
GCSE grades A*-C or equivalent-Degree or equivalent       -2.75203202
Other qualifications-Degree or equivalent                 -1.79074008
No qualification-Degree or equivalent                     -1.82598147
Don't know-Degree or equivalent                           -0.90189294
GCE, A-level or equivalent-Higher education               -0.24615236
GCSE grades A*-C or equivalent-Higher education           -0.58402039
Other qualifications-Higher education                      0.44581610
No qualification-Higher education                          0.39849468
Don't know-Higher education                                1.49495277
GCSE grades A*-C or equivalent-GCE, A-level or equivalent -1.19145627
Other qualifications-GCE, A-level or equivalent           -0.20829388
No qualification-GCE, A-level or equivalent               -0.24717789
Don't know-GCE, A-level or equivalent                      0.72307119
Other qualifications-GCSE grades A*-C or equivalent        0.10307912
No qualification-GCSE grades A*-C or equivalent            0.06354042
Don't know-GCSE grades A*-C or equivalent                  1.04256399
No qualification-Other qualifications                     -1.38996406
Don't know-Other qualifications                           -0.28127508
Don't know-No qualification                               -0.16885517
                                                                  upr
p adj
Higher education-Degree or equivalent                    -1.4725784
0.0000000
GCE, A-level or equivalent-Degree or equivalent          -0.8912923
0.0000000
GCSE grades A*-C or equivalent-Degree or equivalent      -1.1964979
0.0000000
Other qualifications-Degree or equivalent                 0.3208203
0.3814087
No qualification-Degree or equivalent                     0.1701251
0.1796085
Don't know-Degree or equivalent                           4.9840746
0.3861438
GCE, A-level or equivalent-Higher education               1.9511361
0.2497452
GCSE grades A*-C or equivalent-Higher education           1.6400515
0.8020850
Other qualifications-Higher education                     3.0888251
0.0015656
No qualification-Higher education                         2.9502099
0.0021051
```

```
Don't know-Higher education                                    7.5917899
0.0002250
GCSE grades A*-C or equivalent-GCE, A-level or equivalent  0.5425037
0.9272683
Other qualifications-GCE, A-level or equivalent            2.0379514
0.1976841
No qualification-GCE, A-level or equivalent                1.8908988
0.2602847
Don't know-GCE, A-level or equivalent                      6.6586878
0.0046009
Other qualifications-GCSE grades A*-C or equivalent        2.3755309
0.0221266
No qualification-GCSE grades A*-C or equivalent            2.2291331
0.0298123
Don't know-GCSE grades A*-C or equivalent                  6.9881476
0.0013335
No qualification-Other qualifications                      1.2040275
0.9999926
Don't know-Other qualifications                            5.8333766
0.1040286
Don't know-No qualification                                5.9068933
0.0787523



 Hours by Region
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = TOTHRS ~ GOVTOF2, data = qlfs_2015_imputed)

$GOVTOF2
                                                          diff         lwr
upr
North West (inc Merseyside)-North East        0.97828348 -0.72039811
2.6769651
East Midlands-North East                      0.44809329 -1.30962279
2.2058094
West Midlands-North East                      1.59717455 -0.22334408
3.4176932
Eastern-North East                            1.20385180 -0.57114719
2.9788508
London-North East                             0.65435057 -1.08964941
2.3983505
South East-North East                         1.29231381 -0.41549280
3.0001204
South West-North East                         0.34438769 -1.30461157
1.9933870
Wales-North East                              0.25496068 -1.50452090
2.0144423
```

```
Scotland-North East                                1.00936339 -0.95688433
2.9756111
Northern Ireland-North East                        0.59368523 -1.17430321
2.3616737
Does not apply-North East                          2.15760541  0.08567482
4.2295360
East Midlands-North West (inc Merseyside)         -0.53019019 -1.88472862
0.8243482
West Midlands-North West (inc Merseyside)          0.61889107 -0.81620356
2.0539857
Eastern-North West (inc Merseyside)                0.22556832 -1.15132309
1.6024597
London-North West (inc Merseyside)                -0.32393292 -1.66062451
1.0127587
South East-North West (inc Merseyside)             0.31403033 -0.97508252
1.6031432
South West-North West (inc Merseyside)            -0.63389579 -1.84402268
0.5762311
Wales-North West (inc Merseyside)                 -0.72332280 -2.08015144
0.6335058
Scotland-North West (inc Merseyside)               0.03107990 -1.58488377
1.6470436
Northern Ireland-North West (inc Merseyside) -0.38459826 -1.75244022
0.9832437
Does not apply-North West (inc Merseyside)         1.17932193 -0.56369410
2.9223380
West Midlands-East Midlands                        1.14908126 -0.35542638
2.6535889
Eastern-East Midlands                              0.75575851 -0.69333658
2.2048536
London-East Midlands                               0.20625728 -1.20469668
1.6172112
South East-East Midlands                           0.84422052 -0.52174382
2.2101849
South West-East Midlands                          -0.10370560 -1.39539199
1.1879808
Wales-East Midlands                               -0.19313261 -1.62317827
1.2369131
Scotland-East Midlands                             0.56127010 -1.11664088
2.2391811
Northern Ireland-East Midlands                     0.14559194 -1.29490738
1.5860913
Does not apply-East Midlands                       1.70951212 -0.09108543
3.5101097
Eastern-West Midlands                             -0.39332275 -1.91798628
1.1313408
London-West Midlands                              -0.94282399 -2.43128404
0.5456361
South East-West Midlands                          -0.30486074 -1.75074485
1.1410234
```

```
South West-West Midlands                    -1.25278686 -2.62871418
0.1231405
Wales-West Midlands                         -1.34221387 -2.84878377
0.1643560
Scotland-West Midlands                      -0.58781117 -2.33140158
1.1557792
Northern Ireland-West Midlands              -1.00348933 -2.51998547
0.5130068
Does not apply-West Midlands                 0.56043086 -1.30152344
2.4223852
London-Eastern                              -0.54950123 -1.98192808
0.8829256
South East-Eastern                           0.08846201 -1.29967134
1.4765954
South West-Eastern                          -0.85946411 -2.17457222
0.4556440
Wales-Eastern                               -0.94889112 -2.40012721
0.5023450
Scotland-Eastern                            -0.19448841 -1.89049574
1.5015189
Northern Ireland-Eastern                    -0.61016657 -2.07170477
0.8513716
Does not apply-Eastern                       0.95375361 -0.86371912
2.7712263
South East-London                            0.63796325 -0.71030552
1.9862320
South West-London                           -0.30996287 -1.58292156
0.9629958
Wales-London                                -0.39938989 -1.81254263
1.0137629
Scotland-London                              0.35501282 -1.30852418
2.0185498
Northern Ireland-London                     -0.06066534 -1.48439578
1.3630651
Does not apply-London                        1.50325484 -0.28395572
3.2904654
South West-South East                       -0.94792612 -2.17082895
0.2749767
Wales-South East                            -1.03735313 -2.40558857
0.3308823
Scotland-South East                         -0.28295043 -1.90850352
1.3426027
Northern Ireland-South East                 -0.69862859 -2.07778626
0.6805291
Does not apply-South East                    0.86529160 -0.88661854
2.6172017
Wales-South West                            -0.08942701 -1.38351486
1.2046608
Scotland-South West                          0.66497569 -0.89867964
2.2286310
```

```
Northern Ireland-South West                      0.24929753 -1.05633298
1.5549280
Does not apply-South West                        1.81321772  0.11858375
3.5078517
Scotland-Wales                                   0.75440271 -0.92535765
2.4341631
Northern Ireland-Wales                           0.33872455 -1.10392853
1.7813776
Does not apply-Wales                             1.90264473  0.10032369
3.7049658
Northern Ireland-Scotland                       -0.41567816 -2.10434703
1.2729907
Does not apply-Scotland                          1.14824202 -0.85643146
3.1529155
Does not apply-Northern Ireland                  1.56392019 -0.24670646
3.3745468
                                                      p adj
North West (inc Merseyside)-North East       0.7705752
East Midlands-North East                     0.9995846
West Midlands-North East                     0.1527269
Eastern-North East                           0.5373245
London-North East                            0.9869801
South East-North East                        0.3569481
South West-North East                        0.9999414
Wales-North East                             0.9999986
Scotland-North East                          0.8786044
Northern Ireland-North East                  0.9948500
Does not apply-North East                    0.0325528
East Midlands-North West (inc Merseyside)    0.9817573
West Midlands-North West (inc Merseyside)    0.9620008
Eastern-North West (inc Merseyside)          0.9999951
London-North West (inc Merseyside)           0.9997451
South East-North West (inc Merseyside)       0.9997319
South West-North West (inc Merseyside)       0.8631422
Wales-North West (inc Merseyside)            0.8485406
Scotland-North West (inc Merseyside)         1.0000000
Northern Ireland-North West (inc Merseyside) 0.9989525
Does not apply-North West (inc Merseyside)   0.5412377
West Midlands-East Midlands                  0.3421352
Eastern-East Midlands                        0.8666121
London-East Midlands                         0.9999985
South East-East Midlands                     0.6798156
South West-East Midlands                     1.0000000
Wales-East Midlands                          0.9999994
Scotland-East Midlands                       0.9950184
Northern Ireland-East Midlands               1.0000000
Does not apply-East Midlands                 0.0815845
Eastern-West Midlands                        0.9995346
London-West Midlands                         0.6442635
South East-West Midlands                     0.9999355
```

```
South West-West Midlands                    0.1156607
Wales-West Midlands                         0.1365989
Scotland-West Midlands                      0.9946717
Northern Ireland-West Midlands              0.5770142
Does not apply-West Midlands                0.9980432
London-Eastern                              0.9844435
South East-Eastern                          1.0000000
South West-Eastern                          0.5966277
Wales-Eastern                               0.5958646
Scotland-Eastern                            0.9999999
Northern Ireland-Eastern                    0.9700903
Does not apply-Eastern                      0.8616961
South East-London                           0.9275234
South West-London                           0.9997330
Wales-London                                0.9989013
Scotland-London                             0.9999272
Northern Ireland-London                     1.0000000
Does not apply-London                       0.2024629
South West-South East                       0.3188789
Wales-South East                            0.3538575
Scotland-South East                         0.9999908
Northern Ireland-South East                 0.8880151
Does not apply-South East                   0.9043075
Wales-South West                            1.0000000
Scotland-South West                         0.9656944
Northern Ireland-South West                 0.9999763
Does not apply-South West                   0.0238326
Scotland-Wales                              0.9492323
Northern Ireland-Wales                      0.9998128
Does not apply-Wales                        0.0278946
Northern Ireland-Scotland                   0.9997034
Does not apply-Scotland                     0.7768101
Does not apply-Northern Ireland             0.1702037
```

```r
# Create binary employment status (employed vs not employed)
qlfs_2015_imputed$employed_binary <- ifelse(qlfs_2015_imputed$STAT3R
== "Employee", 1, 0)

# Fit multinomial logistic regression
multinom_model <- multinom(STAT3R ~ HIQUL15D + ILODEFR + NSECMJ3R +
GOVTOF2 +
                           SEX + AGEEULR + MARSTA3R + FTPTWK + ETHUK7R,
                           data = qlfs_2015_imputed)

# Display model results
cat("Multinomial Logistic Regression Results\n")
cat("=====================================\n\n")
print(summary(multinom_model))
```

```
# weights:  141 (92 variable)
initial  value 24639.676410
iter  10 value 5395.720973
iter  20 value 5004.082540
iter  30 value 4929.306670
iter  40 value 4891.464499
iter  50 value 4866.279402
iter  60 value 4817.871067
iter  70 value 4790.696704
iter  80 value 4789.410718
iter  90 value 4789.248116
iter 100 value 4788.966873
final  value 4788.966873
stopped after 100 iterations
Multinomial Logistic Regression Results
====================================


Warning message in sqrt(diag(vc)):
“NaNs produced”

Call:
multinom(formula = STAT3R ~ HIQUL15D + ILODEFR + NSECMJ3R + GOVTOF2 +
    SEX + AGEEULR + MARSTA3R + FTPTWK + ETHUK7R, data =
qlfs_2015_imputed)

Coefficients:
                                          (Intercept) HIQUL15DHigher
education
Self-employed                               -3.801193              -
0.01083387
Government scheme or unpaid family worker   -9.213309              -
0.15356717
                                          HIQUL15DGCE, A-level or
equivalent
Self-employed                                                      -
0.1824827
Government scheme or unpaid family worker                          -
0.9483609
                                          HIQUL15DGCSE grades A*-C or
equivalent
Self-employed
-0.3651027
Government scheme or unpaid family worker
-0.3462338
                                          HIQUL15DOther qualifications
Self-employed                                                 0.2617824
Government scheme or unpaid family worker                    -1.2035301
                                          HIQUL15DNo qualification
Self-employed                                                 0.6609854
```

```
Government scheme or unpaid family worker                -0.2745168
                                          HIQUL15DDon't know
Self-employed                                  -0.542447791
Government scheme or unpaid family worker         0.004658493
                                          ILODEFRILO unemployed
ILODEFRInactive
Self-employed                                     0.2037768      -
0.8360348
Government scheme or unpaid family worker            -2.7893366      -
7.2216999
                                          ILODEFRUnder 16
Self-employed                                              0
Government scheme or unpaid family worker                  0
                                          NSECMJ3RIntermediate
occupations and small employers
Self-employed
2.286351
Government scheme or unpaid family worker
0.866469
                                          NSECMJ3RRoutine and manual
occupations
Self-employed
-11.127308
Government scheme or unpaid family worker
1.159201
                                          NSECMJ3RNever worked,
unemployed, and nec
Self-employed
-2.278119
Government scheme or unpaid family worker
5.854308
                                          GOVTOF2North West (inc
Merseyside)
Self-employed
0.1841009
Government scheme or unpaid family worker                        -
0.3750644
                                          GOVTOF2Yorkshire and
Humberside
Self-employed
0
Government scheme or unpaid family worker
0
                                          GOVTOF2East Midlands
Self-employed                                        0.2912652
Government scheme or unpaid family worker           -0.6792913
                                          GOVTOF2West Midlands
GOVTOF2Eastern
Self-employed                                     0.2661120
```

```
0.06611377
Government scheme or unpaid family worker            0.4282181
0.46447883
                                          GOVTOF2London GOVTOF2South
East
Self-employed                                0.3239989
0.7058169
Government scheme or unpaid family worker    0.4584325        -
0.1271910
                                          GOVTOF2South West
GOVTOF2Wales
Self-employed                                 0.3542691
0.4623636
Government scheme or unpaid family worker       -0.3644160   -
0.1144023
                                          GOVTOF2Scotland
Self-employed                                   0.1987268
Government scheme or unpaid family worker       1.3590957
                                          GOVTOF2Northern Ireland
Self-employed                                    0.01961854
Government scheme or unpaid family worker        0.22030485
                                          GOVTOF2Does not apply
SEXFemale
Self-employed                                   0.1846172 -
1.4956039
Government scheme or unpaid family worker        1.4787342 -
0.5897341
                                          AGEEULR20-24 AGEEULR25-29
Self-employed                                0.7063245    0.9187675
Government scheme or unpaid family worker    0.3403238    1.1901009
                                          AGEEULR30-34 AGEEULR35-39
Self-employed                                 1.149493     1.392382
Government scheme or unpaid family worker     1.091015     1.804638
                                          AGEEULR40-44 AGEEULR45-49
Self-employed                                 1.599594    1.8406588
Government scheme or unpaid family worker     1.580353    0.9416497
                                          AGEEULR50-54 AGEEULR55-59
Self-employed                                1.8244132      1.88729
Government scheme or unpaid family worker    0.5536422      2.64565
                                          AGEEULR60-64 AGEEULR65-69
Self-employed                                 1.954340     2.040411
Government scheme or unpaid family worker     2.432752     3.609087
                                          AGEEULR70 and over
Self-employed                                       2.463795
Government scheme or unpaid family worker           3.714419

MARSTA3RMarried/Cohabiting/Civil Partner
Self-employed
0.06217765
```

```
Government scheme or unpaid family worker
0.27147746

MARSTA3RDivorced/Widowed/Previously in Civil Partnership
Self-employed
0.3271480
Government scheme or unpaid family worker
-0.1706729
                                          FTPTWKPart-time
Self-employed                                   0.6550133
Government scheme or unpaid family worker       0.9882054
                                          ETHUK7RMixed/Multiple ethnic
groups
Self-employed                                                    -
0.1300639
Government scheme or unpaid family worker
0.6092969
                                          ETHUK7RIndian
Self-employed                                 -0.182734
Government scheme or unpaid family worker    -17.410493
                                          ETHUK7RPakistani or
Bangladeshi
Self-employed
0.4028021
Government scheme or unpaid family worker
0.3792401
                                          ETHUK7RChinese or any other
Asian background
Self-employed
0.07866494
Government scheme or unpaid family worker
-0.17521578

ETHUK7RBlack/African/Caribbean/Black British
Self-employed
-0.3602945
Government scheme or unpaid family worker
-0.4731302
                                          ETHUK7ROther ethnic group
Self-employed                                          0.1036389
Government scheme or unpaid family worker             -0.3729552

Std. Errors:
                                          (Intercept) HIQUL15DHigher
education
Self-employed                               0.4416828
0.09387072
Government scheme or unpaid family worker   1.0400143
0.45553969
                                          HIQUL15DGCE, A-level or
```

```
equivalent
Self-employed
0.07821069
Government scheme or unpaid family worker
0.42516199
                                          HIQUL15DGCSE grades A*-C or
equivalent
Self-employed
0.08523692
Government scheme or unpaid family worker
0.45765585
                                          HIQUL15DOther qualifications
Self-employed                                                0.1197465
Government scheme or unpaid family worker                    0.8040674
                                          HIQUL15DNo qualification
Self-employed                                 0.1280563
Government scheme or unpaid family worker     0.5581805
                                          HIQUL15DDon't know
Self-employed                                0.3170908
Government scheme or unpaid family worker    1.1699859
                                          ILODEFRILO unemployed
ILODEFRInactive
Self-employed                                   0.2149765
0.09415759
Government scheme or unpaid family worker       0.6645889
0.77687514
                                          ILODEFRUnder 16
Self-employed                              4.611097e-12
Government scheme or unpaid family worker  1.239290e-13
                                          NSECMJ3RIntermediate
occupations and small employers
Self-employed
0.06334484
Government scheme or unpaid family worker
0.77674028
                                          NSECMJ3RRoutine and manual
occupations
Self-employed
10.2891408
Government scheme or unpaid family worker
0.6412256
                                          NSECMJ3RNever worked,
unemployed, and nec
Self-employed
0.1844189
Government scheme or unpaid family worker
0.5661828
                                          GOVTOF2North West (inc
Merseyside)
```

```
Self-employed
0.1719508
Government scheme or unpaid family worker
0.8974867
                                          GOVTOF2Yorkshire and
Humberside
Self-employed
1.719917e-12
Government scheme or unpaid family worker
NaN
                                          GOVTOF2East Midlands
Self-employed                                        0.1768163
Government scheme or unpaid family worker            0.9574795
                                          GOVTOF2West Midlands
GOVTOF2Eastern
Self-employed                                         0.181678
0.1807789
Government scheme or unpaid family worker             0.863637
0.8491378
                                          GOVTOF2London GOVTOF2South
East
Self-employed                                 0.1732553
0.1714884
Government scheme or unpaid family worker     0.8461411
0.8745738
                                          GOVTOF2South West
GOVTOF2Wales
Self-employed                                    0.1647644
0.1729231
Government scheme or unpaid family worker         0.8468609
0.8855244
                                          GOVTOF2Scotland
Self-employed                                   0.1973275
Government scheme or unpaid family worker       0.8501956
                                          GOVTOF2Northern Ireland
Self-employed                                         0.1802165
Government scheme or unpaid family worker             0.8375561
                                          GOVTOF2Does not apply
SEXFemale
Self-employed                                        0.2119551
0.06404382
Government scheme or unpaid family worker            0.8596464
0.27887915
                                          AGEEULR20-24 AGEEULR25-29
Self-employed                                0.4374394    0.4249271
Government scheme or unpaid family worker    0.5500942    0.6474043
                                          AGEEULR30-34 AGEEULR35-39
Self-employed                                0.4230588    0.4231907
Government scheme or unpaid family worker    0.7280880    0.7382280
```

```
                                          AGEEULR40-44 AGEEULR45-49
Self-employed                                0.4208959    0.4208416
Government scheme or unpaid family worker    0.7506198    0.9340488
                                          AGEEULR50-54 AGEEULR55-59
Self-employed                                0.4214091    0.4226295
Government scheme or unpaid family worker    1.2003981    0.7203156
                                          AGEEULR60-64 AGEEULR65-69
Self-employed                                0.4247425    0.4281860
Government scheme or unpaid family worker    0.8362096    0.7516979
                                          AGEEULR70 and over
Self-employed                                      0.4343093
Government scheme or unpaid family worker          0.8392268

MARSTA3RMarried/Cohabiting/Civil Partner
Self-employed
0.07818973
Government scheme or unpaid family worker
0.46270780

MARSTA3RDivorced/Widowed/Previously in Civil Partnership
Self-employed
0.1080334
Government scheme or unpaid family worker
0.6655764
                                          FTPTWKPart-time
Self-employed                                  0.06764067
Government scheme or unpaid family worker      0.33564202
                                          ETHUK7RMixed/Multiple ethnic
groups
Self-employed
0.3085917
Government scheme or unpaid family worker
1.0747785
                                          ETHUK7RIndian
Self-employed                              1.834828e-01
Government scheme or unpaid family worker  2.654419e-08
                                          ETHUK7RPakistani or
Bangladeshi
Self-employed
0.2087947
Government scheme or unpaid family worker
1.0805714
                                          ETHUK7RChinese or any other
Asian background
Self-employed
0.2437014
Government scheme or unpaid family worker
0.9735814

ETHUK7RBlack/African/Caribbean/Black British
```

```
Self-employed
0.2250858
Government scheme or unpaid family worker
0.6910744
                                          ETHUK7ROther ethnic group
Self-employed                                                0.2447526
Government scheme or unpaid family worker                    1.1191998

Residual Deviance: 9577.934
AIC: 9753.934
```

```
# Binary logistic regression
binary_model <- glm(employed_binary ~ HIQUL15D + ILODEFR + NSECMJ3R +
GOVTOF2 +
                      SEX + AGEEULR + MARSTA3R + FTPTWK + ETHUK7R,
                      family = binomial(link = "logit"),
                      data = qlfs_2015_imputed)

# Display binary model results
cat("Binary Logistic Regression Results\n")
cat("==================================\n\n")
print(summary(binary_model))
```

```
Binary Logistic Regression Results
==================================


Call:
glm(formula = employed_binary ~ HIQUL15D + ILODEFR + NSECMJ3R +
    GOVTOF2 + SEX + AGEEULR + MARSTA3R + FTPTWK + ETHUK7R, family =
binomial(link = "logit"),
    data = qlfs_2015_imputed)

Coefficients:
                                                      Estimate Std.
Error
(Intercept)                                            3.24155
0.31571
HIQUL15DHigher education                              -0.01158
0.09226
HIQUL15DGCE, A-level or equivalent                     0.19902
0.07714
HIQUL15DGCSE grades A*-C or equivalent                 0.36023
0.08402
HIQUL15DOther qualifications                          -0.21492
0.11751
HIQUL15DNo qualification                              -0.55463
0.12138
HIQUL15DDon't know                                     0.49294
0.30864
```

```
ILODEFRILO unemployed                                        0.08135
0.19657
ILODEFRInactive                                              1.14559
0.09387
NSECMJ3RIntermediate occupations and small employers        -2.27583
0.06299
NSECMJ3RRoutine and manual occupations                       4.25985
0.33953
NSECMJ3RNever worked, unemployed, and nec                    1.04556
0.13145
GOVTOF2North West (inc Merseyside)                          -0.19498
0.16799
GOVTOF2East Midlands                                        -0.29065
0.17266
GOVTOF2West Midlands                                        -0.31134
0.17724
GOVTOF2Eastern                                              -0.10668
0.17622
GOVTOF2London                                               -0.35579
0.16921
GOVTOF2South East                                           -0.70763
0.16755
GOVTOF2South West                                           -0.35596
0.16111
GOVTOF2Wales                                                -0.46513
0.16895
GOVTOF2Scotland                                             -0.29097
0.19133
GOVTOF2Northern Ireland                                     -0.07485
0.17563
GOVTOF2Does not apply                                       -0.31401
0.20402
SEXFemale                                                    1.49851
0.06234
AGEEULR20-24                                                -0.21619
0.30541
AGEEULR25-29                                                -0.39517
0.29306
AGEEULR30-34                                                -0.57366
0.29060
AGEEULR35-39                                                -0.82736
0.29081
AGEEULR40-44                                                -1.01786
0.28758
AGEEULR45-49                                                -1.24348
0.28743
AGEEULR50-54                                                -1.22577
0.28827
AGEEULR55-59                                                -1.33109
```

```
0.28953
AGEEULR60-64                                                -1.41740
0.29242
AGEEULR65-69                                                -1.60770
0.29558
AGEEULR70 and over                                          -1.79130
0.30062
MARSTA3RMarried/Cohabiting/Civil Partner                    -0.06025
0.07674
MARSTA3RDivorced/Widowed/Previously in Civil Partnership -0.29644
0.10533
FTPTWKPart-time                                             -0.80888
0.06544
ETHUK7RMixed/Multiple ethnic groups                          0.11151
0.30004
ETHUK7RIndian                                                0.20251
0.18088
ETHUK7RPakistani or Bangladeshi                             -0.35639
0.20180
ETHUK7RChinese or any other Asian background                -0.12867
0.23403
ETHUK7RBlack/African/Caribbean/Black British                 0.31258
0.21297
ETHUK7ROther ethnic group                                   -0.08393
0.23731
                                                           z value Pr(>|
z|)
(Intercept)                                                 10.267  <
2e-16 ***
HIQUL15DHigher education                                    -0.125
0.900139
HIQUL15DGCE, A-level or equivalent                           2.580
0.009879 **
HIQUL15DGCSE grades A*-C or equivalent                       4.287
1.81e-05 ***
HIQUL15DOther qualifications                                -1.829
0.067404 .
HIQUL15DNo qualification                                    -4.569
4.89e-06 ***
HIQUL15DDon't know                                           1.597
0.110236
ILODEFRILO unemployed                                        0.414
0.679006
ILODEFRInactive                                             12.204  <
2e-16 ***
NSECMJ3RIntermediate occupations and small employers     -36.130  <
2e-16 ***
NSECMJ3RRoutine and manual occupations                      12.546  <
2e-16 ***
```

```
NSECMJ3RNever worked, unemployed, and nec                    7.954
1.80e-15 ***
GOVTOF2North West (inc Merseyside)                          -1.161
0.245799
GOVTOF2East Midlands                                        -1.683
0.092307 .
GOVTOF2West Midlands                                        -1.757
0.078978 .
GOVTOF2Eastern                                              -0.605
0.544904
GOVTOF2London                                               -2.103
0.035495 *
GOVTOF2South East                                           -4.223
2.41e-05 ***
GOVTOF2South West                                           -2.209
0.027146 *
GOVTOF2Wales                                                -2.753
0.005902 **
GOVTOF2Scotland                                             -1.521
0.128328
GOVTOF2Northern Ireland                                     -0.426
0.669995
GOVTOF2Does not apply                                       -1.539
0.123778
SEXFemale                                                   24.036  <
2e-16 ***
AGEEULR20-24                                                -0.708
0.479037
AGEEULR25-29                                                -1.348
0.177527
AGEEULR30-34                                                -1.974
0.048375 *
AGEEULR35-39                                                -2.845
0.004441 **
AGEEULR40-44                                                -3.539
0.000401 ***
AGEEULR45-49                                                -4.326
1.52e-05 ***
AGEEULR50-54                                                -4.252
2.12e-05 ***
AGEEULR55-59                                                -4.597
4.28e-06 ***
AGEEULR60-64                                                -4.847
1.25e-06 ***
AGEEULR65-69                                                -5.439
5.36e-08 ***
AGEEULR70 and over                                          -5.959
2.54e-09 ***
MARSTA3RMarried/Cohabiting/Civil Partner                    -0.785
```

```
0.432379
MARSTA3RDivorced/Widowed/Previously in Civil Partnership  -2.814
0.004886 **
FTPTWKPart-time                                          -12.361  <
2e-16 ***
ETHUK7RMixed/Multiple ethnic groups                        0.372
0.710155
ETHUK7RIndian                                              1.120
0.262897
ETHUK7RPakistani or Bangladeshi                           -1.766
0.077394 .
ETHUK7RChinese or any other Asian background              -0.550
0.582461
ETHUK7RBlack/African/Caribbean/Black British               1.468
0.142173
ETHUK7ROther ethnic group                                 -0.354
0.723586
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 15538.8  on 22427  degrees of freedom
Residual deviance:  9671.1  on 22384  degrees of freedom
AIC: 9759.1

Number of Fisher Scoring iterations: 9


# Regional Analysis Models (first 3 regions for brevity)
regions <- unique(qlfs_2015_imputed$GOVTOF2)[1:3]
regional_models <- lapply(regions, function(region) {
  region_data <- subset(qlfs_2015_imputed, GOVTOF2 == region)
  glm(employed_binary ~ HIQUL15D + ILODEFR + NSECMJ3R +
      SEX + AGEEULR + MARSTA3R + FTPTWK + ETHUK7R,
      family = binomial(link = "logit"),
      data = region_data)
})

# Display regional model results
for (i in seq_along(regional_models)) {
  cat("\nRegion:", regions[i], "\n")
  print(summary(regional_models[[i]]))
  cat("\n")
}

Warning message:
“glm.fit: fitted probabilities numerically 0 or 1 occurred”


Region: 9
```

```
Call:
glm(formula = employed_binary ~ HIQUL15D + ILODEFR + NSECMJ3R +
    SEX + AGEEULR + MARSTA3R + FTPTWK + ETHUK7R, family =
binomial(link = "logit"),
    data = region_data)

Coefficients:
                                                      Estimate
Std. Error
(Intercept)                                           2.876150
0.607009
HIQUL15DHigher education                             -0.006861
0.225593
HIQUL15DGCE, A-level or equivalent                    0.322532
0.193275
HIQUL15DGCSE grades A*-C or equivalent                0.216765
0.203331
HIQUL15DOther qualifications                         -0.169507
0.314593
HIQUL15DNo qualification                             -0.683442
0.324980
HIQUL15DDon't know                                   -0.942932
1.114431
ILODEFRILO unemployed                                 0.603535
0.604851
ILODEFRInactive                                       1.139650
0.228836
NSECMJ3RIntermediate occupations and small employers -2.185521
0.157410
NSECMJ3RRoutine and manual occupations                4.554237
1.016485
NSECMJ3RNever worked, unemployed, and nec             1.057820
0.356018
SEXFemale                                             1.521832
0.162788
AGEEULR20-24                                          0.999567
0.933929
AGEEULR25-29                                         -0.637465
0.651146
AGEEULR30-34                                         -0.840020
0.652652
AGEEULR35-39                                         -1.207857
0.653063
AGEEULR40-44                                         -0.771475
0.645732
AGEEULR45-49                                         -1.443335
0.633270
AGEEULR50-54                                         -1.495856
```

```
0.636928
AGEEULR55-59                                                  -1.503231
0.640201
AGEEULR60-64                                                  -1.629528
0.646718
AGEEULR65-69                                                  -2.179103
0.649605
AGEEULR70 and over                                            -1.897651
0.665684
MARSTA3RMarried/Cohabiting/Civil Partner                       0.222890
0.205735
MARSTA3RDivorced/Widowed/Previously in Civil Partnership      -0.224206
0.266827
FTPTWKPart-time                                               -1.012729
0.161870
ETHUK7RMixed/Multiple ethnic groups                           0.305821
0.743398
ETHUK7RIndian                                                 0.818377
0.693987
ETHUK7RPakistani or Bangladeshi                               -0.027781
0.704058
ETHUK7RChinese or any other Asian background                  -0.149958
0.566669
ETHUK7RBlack/African/Caribbean/Black British                  -0.753207
0.653616
ETHUK7ROther ethnic group                                     0.645371
0.671475
                                                              z value Pr(>|
z|)
(Intercept)                                                     4.738
2.16e-06 ***
HIQUL15DHigher education                                      -0.030
0.975739
HIQUL15DGCE, A-level or equivalent                            1.669
0.095161 .
HIQUL15DGCSE grades A*-C or equivalent                        1.066
0.286392
HIQUL15DOther qualifications                                  -0.539
0.590016
HIQUL15DNo qualification                                      -2.103
0.035463 *
HIQUL15DDon't know                                            -0.846
0.397491
ILODEFRILO unemployed                                         0.998
0.318364
ILODEFRInactive                                               4.980
6.35e-07 ***
NSECMJ3RIntermediate occupations and small employers          -13.884  <
2e-16 ***
```

```
NSECMJ3RRoutine and manual occupations                        4.480
7.45e-06 ***
NSECMJ3RNever worked, unemployed, and nec                     2.971
0.002966 **
SEXFemale                                                     9.349  <
2e-16 ***
AGEEULR20-24                                                  1.070
0.284493
AGEEULR25-29                                                 -0.979
0.327585
AGEEULR30-34                                                 -1.287
0.198064
AGEEULR35-39                                                 -1.850
0.064382 .
AGEEULR40-44                                                 -1.195
0.232193
AGEEULR45-49                                                 -2.279
0.022656 *
AGEEULR50-54                                                 -2.349
0.018847 *
AGEEULR55-59                                                 -2.348
0.018871 *
AGEEULR60-64                                                 -2.520
0.011746 *
AGEEULR65-69                                                 -3.355
0.000795 ***
AGEEULR70 and over                                           -2.851
0.004363 **
MARSTA3RMarried/Cohabiting/Civil Partner                      1.083
0.278638
MARSTA3RDivorced/Widowed/Previously in Civil Partnership     -0.840
0.400760
FTPTWKPart-time                                              -6.256
3.94e-10 ***
ETHUK7RMixed/Multiple ethnic groups                           0.411
0.680792
ETHUK7RIndian                                                 1.179
0.238303
ETHUK7RPakistani or Bangladeshi                              -0.039
0.968524
ETHUK7RChinese or any other Asian background                 -0.265
0.791294
ETHUK7RBlack/African/Caribbean/Black British                 -1.152
0.249169
ETHUK7ROther ethnic group                                     0.961
0.336490
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2364.7  on 3112  degrees of freedom
Residual deviance: 1502.4  on 3080  degrees of freedom
AIC: 1568.4

Number of Fisher Scoring iterations: 9



Region: 10

Call:
glm(formula = employed_binary ~ HIQUL15D + ILODEFR + NSECMJ3R +
    SEX + AGEEULR + MARSTA3R + FTPTWK + ETHUK7R, family =
binomial(link = "logit"),
    data = region_data)

Coefficients:
                                                          Estimate
Std. Error
(Intercept)                                              2.837e+00
1.138e+00
HIQUL15DHigher education                                 1.529e-01
3.070e-01
HIQUL15DGCE, A-level or equivalent                       1.427e-01
2.511e-01
HIQUL15DGCSE grades A*-C or equivalent                   7.720e-01
2.925e-01
HIQUL15DOther qualifications                            -8.228e-01
4.207e-01
HIQUL15DNo qualification                                -7.820e-01
4.271e-01
HIQUL15DDon't know                                      -3.563e-03
1.246e+00
ILODEFRILO unemployed                                    3.581e-01
7.446e-01
ILODEFRInactive                                          1.560e+00
3.142e-01
NSECMJ3RIntermediate occupations and small employers    -2.430e+00
2.101e-01
NSECMJ3RRoutine and manual occupations                   1.760e+01
4.508e+02
NSECMJ3RNever worked, unemployed, and nec                1.997e+00
4.836e-01
SEXFemale                                                1.414e+00
2.095e-01
AGEEULR20-24                                             1.099e-01
1.261e+00
AGEEULR25-29                                            -4.802e-01
```

```
1.176e+00
AGEEULR30-34                                                  -1.437e+00
1.168e+00
AGEEULR35-39                                                  -7.656e-01
1.201e+00
AGEEULR40-44                                                  -1.315e+00
1.172e+00
AGEEULR45-49                                                  -1.548e+00
1.166e+00
AGEEULR50-54                                                  -1.466e+00
1.168e+00
AGEEULR55-59                                                  -1.414e+00
1.173e+00
AGEEULR60-64                                                  -1.863e+00
1.178e+00
AGEEULR65-69                                                  -2.179e+00
1.187e+00
AGEEULR70 and over                                            -3.482e+00
1.214e+00
MARSTA3RMarried/Cohabiting/Civil Partner                       1.918e-01
2.700e-01
MARSTA3RDivorced/Widowed/Previously in Civil Partnership -5.295e-02
3.501e-01
FTPTWKPart-time                                               -5.286e-01
2.177e-01
ETHUK7RMixed/Multiple ethnic groups                            1.721e+01
2.406e+03
ETHUK7RIndian                                                  6.996e-01
1.177e+00
ETHUK7RPakistani or Bangladeshi                                1.587e+01
7.010e+03
ETHUK7RChinese or any other Asian background                  -1.023e+00
1.151e+00
ETHUK7RBlack/African/Caribbean/Black British                   1.787e+01
1.940e+03
ETHUK7ROther ethnic group                                      2.527e-01
9.499e-01
                                                              z value Pr(>|
z|)
(Intercept)                                                     2.493
0.01265 *
HIQUL15DHigher education                                        0.498
0.61853
HIQUL15DGCE, A-level or equivalent                              0.569
0.56969
HIQUL15DGCSE grades A*-C or equivalent                          2.639
0.00832 **
HIQUL15DOther qualifications                                   -1.956
0.05049 .
```

```
HIQUL15DNo qualification                                    -1.831
0.06710 .
HIQUL15DDon't know                                         -0.003
0.99772
ILODEFRILO unemployed                                       0.481
0.63057
ILODEFRInactive                                             4.964
6.91e-07 ***
NSECMJ3RIntermediate occupations and small employers      -11.568  <
2e-16 ***
NSECMJ3RRoutine and manual occupations                      0.039
0.96885
NSECMJ3RNever worked, unemployed, and nec                   4.129
3.65e-05 ***
SEXFemale                                                   6.749
1.49e-11 ***
AGEEULR20-24                                                0.087
0.93050
AGEEULR25-29                                               -0.408
0.68316
AGEEULR30-34                                               -1.230
0.21853
AGEEULR35-39                                               -0.637
0.52389
AGEEULR40-44                                               -1.122
0.26185
AGEEULR45-49                                               -1.328
0.18431
AGEEULR50-54                                               -1.256
0.20920
AGEEULR55-59                                               -1.206
0.22780
AGEEULR60-64                                               -1.582
0.11373
AGEEULR65-69                                               -1.835
0.06648 .
AGEEULR70 and over                                         -2.869
0.00412 **
MARSTA3RMarried/Cohabiting/Civil Partner                    0.710
0.47742
MARSTA3RDivorced/Widowed/Previously in Civil Partnership   -0.151
0.87978
FTPTWKPart-time                                            -2.428
0.01517 *
ETHUK7RMixed/Multiple ethnic groups                         0.007
0.99429
ETHUK7RIndian                                               0.595
0.55214
ETHUK7RPakistani or Bangladeshi                             0.002
```

```
0.99819
ETHUK7RChinese or any other Asian background               -0.888
0.37428
ETHUK7RBlack/African/Caribbean/Black British                0.009
0.99265
ETHUK7ROther ethnic group                                   0.266
0.79019
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1548.38  on 1969  degrees of freedom
Residual deviance:  861.94  on 1937  degrees of freedom
AIC: 927.94

Number of Fisher Scoring iterations: 18



Region: 5

Call:
glm(formula = employed_binary ~ HIQUL15D + ILODEFR + NSECMJ3R +
    SEX + AGEEULR + MARSTA3R + FTPTWK + ETHUK7R, family =
binomial(link = "logit"),
    data = region_data)

Coefficients:
                                                          Estimate Std.
Error
(Intercept)                                                3.8789
1.1073
HIQUL15DHigher education                                  -0.1381
0.3557
HIQUL15DGCE, A-level or equivalent                         0.4710
0.3172
HIQUL15DGCSE grades A*-C or equivalent                     0.3970
0.3421
HIQUL15DOther qualifications                               0.2334
0.4596
HIQUL15DNo qualification                                  -0.3591
0.5718
HIQUL15DDon't know                                        -0.5465
0.8470
ILODEFRILO unemployed                                     -0.3498
0.6872
ILODEFRInactive                                            1.0584
0.3672
NSECMJ3RIntermediate occupations and small employers      -2.5718
```

```
0.2601
NSECMJ3RRoutine and manual occupations                        3.7761
1.0352
NSECMJ3RNever worked, unemployed, and nec                     0.4196
0.4885
SEXFemale                                                     1.7036
0.2475
AGEEULR20-24                                                  0.7211
1.3420
AGEEULR25-29                                                 -1.1564
1.1804
AGEEULR30-34                                                 -1.8657
1.1451
AGEEULR35-39                                                 -1.6504
1.1630
AGEEULR40-44                                                 -1.9804
1.1277
AGEEULR45-49                                                 -1.8720
1.1555
AGEEULR50-54                                                 -2.7315
1.1316
AGEEULR55-59                                                 -1.7467
1.1657
AGEEULR60-64                                                 -2.3417
1.1555
AGEEULR65-69                                                 -2.4090
1.1809
AGEEULR70 and over                                           -2.3816
1.2088
MARSTA3RMarried/Cohabiting/Civil Partner                     -0.1027
0.3145
MARSTA3RDivorced/Widowed/Previously in Civil Partnership   0.1439
0.4159
FTPTWKPart-time                                              -0.8168
0.2584
ETHUK7RMixed/Multiple ethnic groups                          -0.8898
1.3148
ETHUK7RIndian                                                -0.5586
0.5005
ETHUK7RPakistani or Bangladeshi                              -0.1929
1.3065
ETHUK7RChinese or any other Asian background                 -2.0494
2.2102
ETHUK7RBlack/African/Caribbean/Black British                 -0.3374
0.9496
ETHUK7ROther ethnic group                                    -0.1086
1.3142
                                                          z value Pr(>|
z|)
```

```
(Intercept)                                                 3.503
0.000460 ***
HIQUL15DHigher education                                   -0.388
0.697797
HIQUL15DGCE, A-level or equivalent                          1.485
0.137568
HIQUL15DGCSE grades A*-C or equivalent                      1.161
0.245796
HIQUL15DOther qualifications                                0.508
0.611638
HIQUL15DNo qualification                                   -0.628
0.529984
HIQUL15DDon't know                                         -0.645
0.518786
ILODEFRILO unemployed                                      -0.509
0.610762
ILODEFRInactive                                             2.883
0.003945 **
NSECMJ3RIntermediate occupations and small employers       -9.888  <
2e-16 ***
NSECMJ3RRoutine and manual occupations                      3.648
0.000264 ***
NSECMJ3RNever worked, unemployed, and nec                   0.859
0.390321
SEXFemale                                                   6.884
5.81e-12 ***
AGEEULR20-24                                                0.537
0.591050
AGEEULR25-29                                               -0.980
0.327252
AGEEULR30-34                                               -1.629
0.103254
AGEEULR35-39                                               -1.419
0.155866
AGEEULR40-44                                               -1.756
0.079078 .
AGEEULR45-49                                               -1.620
0.105213
AGEEULR50-54                                               -2.414
0.015790 *
AGEEULR55-59                                               -1.498
0.134025
AGEEULR60-64                                               -2.027
0.042705 *
AGEEULR65-69                                               -2.040
0.041361 *
AGEEULR70 and over                                         -1.970
0.048813 *
MARSTA3RMarried/Cohabiting/Civil Partner                   -0.327
```

```
0.744028
MARSTA3RDivorced/Widowed/Previously in Civil Partnership   0.346
0.729376
FTPTWKPart-time                                           -3.160
0.001575 **
ETHUK7RMixed/Multiple ethnic groups                       -0.677
0.498588
ETHUK7RIndian                                             -1.116
0.264377
ETHUK7RPakistani or Bangladeshi                           -0.148
0.882643
ETHUK7RChinese or any other Asian background              -0.927
0.353796
ETHUK7RBlack/African/Caribbean/Black British              -0.355
0.722320
ETHUK7ROther ethnic group                                 -0.083
0.934112
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1084.6  on 1623  degrees of freedom
Residual deviance:  649.5  on 1591  degrees of freedom
AIC: 715.5

Number of Fisher Scoring iterations: 8



# Multicollinearity Check
vif_results <- vif(binary_model)
print("Variance Inflation Factors (VIF):")
print(vif_results)

[1] "Variance Inflation Factors (VIF):"
             GVIF Df GVIF^(1/(2*Df))
HIQUL15D 1.514554  6        1.035199
ILODEFR  1.929850  2        1.178639
NSECMJ3R 1.888673  3        1.111799
GOVTOF2  1.209759 11        1.008693
SEX      1.373745  1        1.172069
AGEEULR  3.418563 11        1.057464
MARSTA3R 1.575616  2        1.120373
FTPTWK   1.309898  1        1.144508
ETHUK7R  1.203405  6        1.015549

# Heteroskedasticity Test
hetero_test <- bptest(binary_model)
cat("Heteroskedasticity Test\n")
```

```r
cat("======================\n\n")
print(hetero_test)
```

```
Heteroskedasticity Test
======================


	studentized Breusch-Pagan test

data:  binary_model
BP = 5658.4, df = 43, p-value < 2.2e-16
```

```r
# Robust Standard Errors
robust_se <- sqrt(diag(vcovHC(binary_model, type = "HC1")))
coef_table <- cbind(
  Estimate = coef(binary_model),
  "Std. Error" = robust_se,
  "z value" = coef(binary_model) / robust_se,
  "Pr(>|z|)" = 2 * pnorm(abs(coef(binary_model) / robust_se),
lower.tail = FALSE)
)
print("Coefficients with Robust Standard Errors:")
print(coef_table)
```

```
[1] "Coefficients with Robust Standard Errors:"
                                                                 Estimate
Std. Error
(Intercept)                                                   3.24154900
0.32285712
HIQUL15DHigher education                                     -0.01157751
0.09455844
HIQUL15DGCE, A-level or equivalent                            0.19901506
0.07632969
HIQUL15DGCSE grades A*-C or equivalent                        0.36022869
0.08523488
HIQUL15DOther qualifications                                 -0.21491906
0.11296256
HIQUL15DNo qualification                                     -0.55462860
0.11552292
HIQUL15DDon't know                                            0.49293851
0.28643568
ILODEFRILO unemployed                                         0.08134564
0.22359175
ILODEFRInactive                                               1.14558970
0.11055144
NSECMJ3RIntermediate occupations and small employers         -2.27582836
0.06437699
NSECMJ3RRoutine and manual occupations                        4.25985201
0.33985637
```

```
NSECMJ3RNever worked, unemployed, and nec                 1.04556017
0.16124688
GOVTOF2North West (inc Merseyside)                       -0.19497532
0.17187933
GOVTOF2East Midlands                                     -0.29065004
0.17372029
GOVTOF2West Midlands                                     -0.31134224
0.18083175
GOVTOF2Eastern                                           -0.10668383
0.18122724
GOVTOF2London                                            -0.35578872
0.17235429
GOVTOF2South East                                        -0.70763146
0.17394871
GOVTOF2South West                                        -0.35595899
0.16529048
GOVTOF2Wales                                             -0.46513293
0.17190787
GOVTOF2Scotland                                          -0.29096816
0.19207522
GOVTOF2Northern Ireland                                  -0.07484736
0.18052653
GOVTOF2Does not apply                                    -0.31401033
0.20402656
SEXFemale                                                 1.49851472
0.06683514
AGEEULR20-24                                             -0.21618545
0.31414065
AGEEULR25-29                                             -0.39516558
0.30305644
AGEEULR30-34                                             -0.57365860
0.29767587
AGEEULR35-39                                             -0.82735630
0.29998646
AGEEULR40-44                                             -1.01786088
0.29454199
AGEEULR45-49                                             -1.24347720
0.29309746
AGEEULR50-54                                             -1.22576752
0.29413975
AGEEULR55-59                                             -1.33108548
0.29675999
AGEEULR60-64                                             -1.41739935
0.29954363
AGEEULR65-69                                             -1.60769665
0.30489368
AGEEULR70 and over                                       -1.79129588
0.30289693
MARSTA3RMarried/Cohabiting/Civil Partner                 -0.06024936
```

```
0.07561874
MARSTA3RDivorced/Widowed/Previously in Civil Partnership -0.29644415
0.10421732
FTPTWKPart-time                                          -0.80887612
0.07177854
ETHUK7RMixed/Multiple ethnic groups                       0.11150806
0.31319484
ETHUK7RIndian                                             0.20250688
0.18987124
ETHUK7RPakistani or Bangladeshi                          -0.35638515
0.20399309
ETHUK7RChinese or any other Asian background             -0.12866807
0.21102347
ETHUK7RBlack/African/Caribbean/Black British              0.31258491
0.22484798
ETHUK7ROther ethnic group                                -0.08392914
0.26038764
                                                            z value
(Intercept)                                              10.0401966
HIQUL15DHigher education                                 -0.1224376
HIQUL15DGCE, A-level or equivalent                        2.6073088
HIQUL15DGCSE grades A*-C or equivalent                    4.2263063
HIQUL15DOther qualifications                             -1.9025689
HIQUL15DNo qualification                                 -4.8010266
HIQUL15DDon't know                                        1.7209396
ILODEFRILO unemployed                                     0.3638132
ILODEFRInactive                                          10.3625038
NSECMJ3RIntermediate occupations and small employers    -35.3515823
NSECMJ3RRoutine and manual occupations                   12.5342714
NSECMJ3RNever worked, unemployed, and nec                 6.4842196
GOVTOF2North West (inc Merseyside)                       -1.1343733
GOVTOF2East Midlands                                     -1.6730921
GOVTOF2West Midlands                                     -1.7217233
GOVTOF2Eastern                                           -0.5886743
GOVTOF2London                                            -2.0642870
GOVTOF2South East                                        -4.0680465
GOVTOF2South West                                        -2.1535359
GOVTOF2Wales                                             -2.7057105
GOVTOF2Scotland                                          -1.5148657
GOVTOF2Northern Ireland                                  -0.4146059
GOVTOF2Does not apply                                    -1.5390660
SEXFemale                                                22.4210593
AGEEULR20-24                                             -0.6881805
AGEEULR25-29                                             -1.3039339
AGEEULR30-34                                             -1.9271250
AGEEULR35-39                                             -2.7579788
AGEEULR40-44                                             -3.4557412
AGEEULR45-49                                             -4.2425383
AGEEULR50-54                                             -4.1672964
```

```
AGEEULR55-59                                                -4.4853940
AGEEULR60-64                                                -4.7318627
AGEEULR65-69                                                -5.2729747
AGEEULR70 and over                                          -5.9138792
MARSTA3RMarried/Cohabiting/Civil Partner                    -0.7967517
MARSTA3RDivorced/Widowed/Previously in Civil Partnership    -2.8444806
FTPTWKPart-time                                            -11.2690516
ETHUK7RMixed/Multiple ethnic groups                          0.3560341
ETHUK7RIndian                                                1.0665485
ETHUK7RPakistani or Bangladeshi                             -1.7470452
ETHUK7RChinese or any other Asian background                -0.6097335
ETHUK7RBlack/African/Caribbean/Black British                 1.3902056
ETHUK7ROther ethnic group                                   -0.3223238
                                                               Pr(>|z|)
(Intercept)                                                1.014717e-23
HIQUL15DHigher education                                   9.025525e-01
HIQUL15DGCE, A-level or equivalent                         9.125703e-03
HIQUL15DGCSE grades A*-C or equivalent                     2.375586e-05
HIQUL15DOther qualifications                               5.709681e-02
HIQUL15DNo qualification                                   1.578543e-06
HIQUL15DDon't know                                         8.526178e-02
ILODEFRILO unemployed                                      7.159975e-01
ILODEFRInactive                                            3.672213e-25
NSECMJ3RIntermediate occupations and small employers      9.480133e-274
NSECMJ3RRoutine and manual occupations                     4.847966e-36
NSECMJ3RNever worked, unemployed, and nec                  8.919221e-11
GOVTOF2North West (inc Merseyside)                         2.566380e-01
GOVTOF2East Midlands                                       9.430917e-02
GOVTOF2West Midlands                                       8.511965e-02
GOVTOF2Eastern                                             5.560798e-01
GOVTOF2London                                              3.899051e-02
GOVTOF2South East                                          4.740892e-05
GOVTOF2South West                                          3.127659e-02
GOVTOF2Wales                                               6.815844e-03
GOVTOF2Scotland                                            1.298064e-01
GOVTOF2Northern Ireland                                    6.784304e-01
GOVTOF2Does not apply                                      1.237882e-01
SEXFemale                                                 2.452785e-111
AGEEULR20-24                                               4.913392e-01
AGEEULR25-29                                               1.922561e-01
AGEEULR30-34                                               5.396406e-02
AGEEULR35-39                                               5.815996e-03
AGEEULR40-44                                               5.487820e-04
AGEEULR45-49                                               2.210058e-05
AGEEULR50-54                                               3.082336e-05
AGEEULR55-59                                               7.277931e-06
AGEEULR60-64                                               2.224688e-06
AGEEULR65-69                                               1.342301e-07
AGEEULR70 and over                                         3.341430e-09
```

```
MARSTA3RMarried/Cohabiting/Civil Partner                    4.255953e-01
MARSTA3RDivorced/Widowed/Previously in Civil Partnership  4.448390e-03
FTPTWKPart-time                                            1.865686e-29
ETHUK7RMixed/Multiple ethnic groups                        7.218150e-01
ETHUK7RIndian                                              2.861758e-01
ETHUK7RPakistani or Bangladeshi                            8.062949e-02
ETHUK7RChinese or any other Asian background               5.420384e-01
ETHUK7RBlack/African/Caribbean/Black British               1.644665e-01
ETHUK7ROther ethnic group                                  7.472074e-01
```

```r
# Calculate feature importance using odds ratios
odds_ratios <- exp(coef(binary_model))
odds_ratios_df <- data.frame(
  Variable = names(odds_ratios),
  Odds_Ratio = odds_ratios
)
print("Odds Ratios (Feature Importance):")
print(odds_ratios_df)

# Create feature importance plot
ggplot(odds_ratios_df, aes(x = reorder(Variable, Odds_Ratio), y =
Odds_Ratio)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  coord_flip() +
  theme_minimal() +
  labs(title = "Feature Importance (Odds Ratios)",
       x = "Variable",
       y = "Odds Ratio")
```

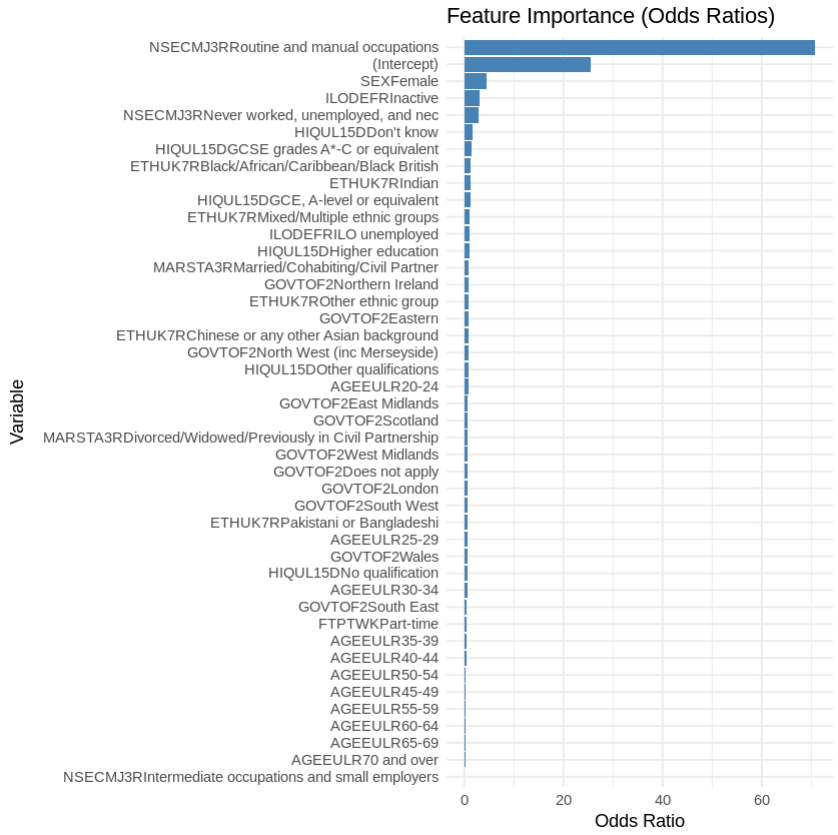
```
[1] "Odds Ratios (Feature Importance):"

Variable
(Intercept)
(Intercept)
HIQUL15DHigher education
HIQUL15DHigher education
HIQUL15DGCE, A-level or equivalent
HIQUL15DGCE, A-level or equivalent
HIQUL15DGCSE grades A*-C or equivalent
HIQUL15DGCSE grades A*-C or equivalent
HIQUL15DOther qualifications
HIQUL15DOther qualifications
HIQUL15DNo qualification
HIQUL15DNo qualification
HIQUL15DDon't know
HIQUL15DDon't know
ILODEFRILO unemployed
ILODEFRILO unemployed
ILODEFRInactive
ILODEFRInactive
```

NSECMJ3RIntermediate occupations and small employers
NSECMJ3RIntermediate occupations and small employers
NSECMJ3RRoutine and manual occupations
NSECMJ3RRoutine and manual occupations
NSECMJ3RNever worked, unemployed, and nec
NSECMJ3RNever worked, unemployed, and nec
GOVTOF2North West (inc Merseyside)
GOVTOF2North West (inc Merseyside)
GOVTOF2East Midlands
GOVTOF2East Midlands
GOVTOF2West Midlands
GOVTOF2West Midlands
GOVTOF2Eastern
GOVTOF2Eastern
GOVTOF2London
GOVTOF2London
GOVTOF2South East
GOVTOF2South East
GOVTOF2South West
GOVTOF2South West
GOVTOF2Wales
GOVTOF2Wales
GOVTOF2Scotland
GOVTOF2Scotland
GOVTOF2Northern Ireland
GOVTOF2Northern Ireland
GOVTOF2Does not apply
GOVTOF2Does not apply
SEXFemale
SEXFemale
AGEEULR20-24
AGEEULR20-24
AGEEULR25-29
AGEEULR25-29
AGEEULR30-34
AGEEULR30-34
AGEEULR35-39
AGEEULR35-39
AGEEULR40-44
AGEEULR40-44
AGEEULR45-49
AGEEULR45-49
AGEEULR50-54
AGEEULR50-54
AGEEULR55-59
AGEEULR55-59
AGEEULR60-64
AGEEULR60-64
AGEEULR65-69

AGEEULR65-69
AGEEULR70 and over
AGEEULR70 and over
MARSTA3RMarried/Cohabiting/Civil Partner
MARSTA3RMarried/Cohabiting/Civil Partner
MARSTA3RDivorced/Widowed/Previously in Civil Partnership
MARSTA3RDivorced/Widowed/Previously in Civil Partnership
FTPTWKPart-time
FTPTWKPart-time
ETHUK7RMixed/Multiple ethnic groups
ETHUK7RMixed/Multiple ethnic groups
ETHUK7RIndian
ETHUK7RIndian
ETHUK7RPakistani or Bangladeshi
ETHUK7RPakistani or Bangladeshi
ETHUK7RChinese or any other Asian background
ETHUK7RChinese or any other Asian background
ETHUK7RBlack/African/Caribbean/Black British
ETHUK7RBlack/African/Caribbean/Black British
ETHUK7ROther ethnic group
ETHUK7ROther ethnic group

	Odds_Ratio
(Intercept)	25.5733042
HIQUL15DHigher education	0.9884893
HIQUL15DGCE, A-level or equivalent	1.2202003
HIQUL15DGCSE grades A*-C or equivalent	1.4336572
HIQUL15DOther qualifications	0.8066067
HIQUL15DNo qualification	0.5742855
HIQUL15DDon't know	1.6371199
ILODEFRILO unemployed	1.0847458
ILODEFRInactive	3.1442950
NSECMJ3RIntermediate occupations and small employers	0.1027118
NSECMJ3RRoutine and manual occupations	70.7995052
NSECMJ3RNever worked, unemployed, and nec	2.8449918
GOVTOF2North West (inc Merseyside)	0.8228550
GOVTOF2East Midlands	0.7477773
GOVTOF2West Midlands	0.7324632
GOVTOF2Eastern	0.8988098
GOVTOF2London	0.7006206
GOVTOF2South East	0.4928101
GOVTOF2South West	0.7005013
GOVTOF2Wales	0.6280516
GOVTOF2Scotland	0.7475395
GOVTOF2Northern Ireland	0.9278851
GOVTOF2Does not apply	0.7305115
SEXFemale	4.4750374
AGEEULR20-24	0.8055859
AGEEULR25-29	0.6735685
AGEEULR30-34	0.5634602

```
AGEEULR35-39                                                  0.4372036
AGEEULR40-44                                                  0.3613671
AGEEULR45-49                                                  0.2883797
AGEEULR50-54                                                  0.2935323
AGEEULR55-59                                                  0.2641903
AGEEULR60-64                                                  0.2423434
AGEEULR65-69                                                  0.2003486
AGEEULR70 and over                                            0.1667439
MARSTA3RMarried/Cohabiting/Civil Partner                      0.9415297
MARSTA3RDivorced/Widowed/Previously in Civil Partnership      0.7434571
FTPTWKPart-time                                               0.4453583
ETHUK7RMixed/Multiple ethnic groups                           1.1179628
ETHUK7RIndian                                                 1.2244685
ETHUK7RPakistani or Bangladeshi                               0.7002029
ETHUK7RChinese or any other Asian background                  0.8792658
ETHUK7RBlack/African/Caribbean/Black British                  1.3669540
ETHUK7ROther ethnic group                                     0.9194964
```



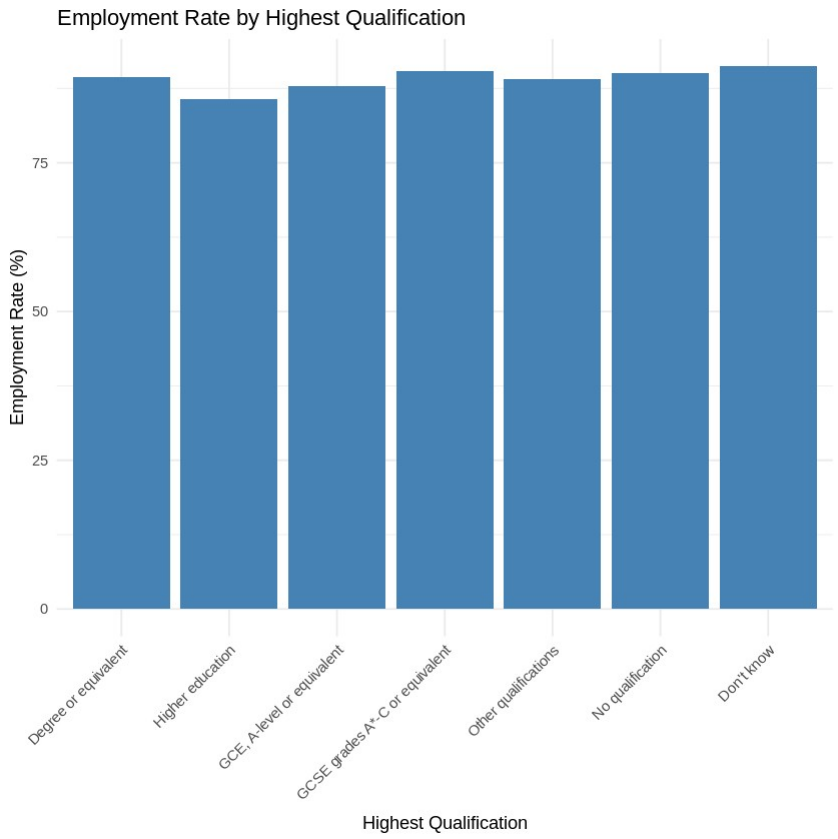

```r
# Calculate AIC and BIC for model comparison
model_comparison <- data.frame(
  Model = c("Binary Logistic", "Multinomial Logistic"),
  AIC = c(AIC(binary_model), AIC(multinom_model)),
  BIC = c(BIC(binary_model), BIC(multinom_model))
```

```
)
print("Model Comparison:")
print(model_comparison)
```

```
[1] "Model Comparison:"
                  Model      AIC      BIC
1      Binary Logistic 9759.060 10111.85
2 Multinomial Logistic 9753.934 10459.52
```

```
# Calculate employment rates
employment_rates <- qlfs_2015_imputed %>%
  group_by(HIQUL15D) %>%
  summarise(
    total = n(),
    employed = sum(employed_binary),
    employment_rate = (employed / total) * 100
  )

# Display employment rates
print("Employment Rates by Qualification:")
print(employment_rates)

# Create employment rate plot
ggplot(employment_rates, aes(x = HIQUL15D, y = employment_rate)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title = "Employment Rate by Highest Qualification",
       x = "Highest Qualification",
       y = "Employment Rate (%)")
```

```
[1] "Employment Rates by Qualification:"
# A tibble: 7 × 4
  HIQUL15D                       total employed employment_rate
  <fct>                          <int>    <dbl>           <dbl>
1 Degree or equivalent            7647     6836            89.4
2 Higher education                1937     1660            85.7
3 GCE, A-level or equivalent      4539     3987            87.8
4 GCSE grades A*-C or equivalent  4195     3793            90.4
5 Other qualifications            1820     1620            89.0
6 No qualification                2096     1889            90.1
7 Don't know                       194      177            91.2
```


Employment Rate by Highest Qualification
Employment Rate (%)
75
50
25
0
Degree or equivalent
Higher education
GCE, A-level or equivalent
GCSE grades A*-C or equivalent
Other qualifications
No qualification
Don't know
Highest Qualification