

# Navigating Factors Affecting Job Satisfaction Among Working Americans

Shirley Chen, Jessica Im

March 4, 2024

## Table of contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Data</b>	<b>2</b>
2.1	Source Data and Methodology . . . . .	2
2.2	Data Cleaning . . . . .	2
2.3	Data Terminology . . . . .	3
2.4	Respondent Demographics . . . . .	3
2.5	Graphs of Responses . . . . .	4
2.5.1	Helps Others . . . . .	4
2.5.2	Interesting Work . . . . .	4
2.5.3	Social Usefulness . . . . .	4
<b>3</b>	<b>Results</b>	<b>6</b>
<b>4</b>	<b>Discussion</b>	<b>10</b>
4.1	Gender? . . . . .	10
4.2	Culture? . . . . .	10
<b>5</b>	<b>Sources</b>	<b>10</b>

# 1 Introduction

## 2 Data

Based at the University of Chicago since 1972, the General Social Survey (GSS) is a project with the objective of monitoring and analyzing the intricacies of American society (1). The GSS Data Explorer makes it so that data retrieved from the project is a publicly available resource, accessible to various types of people, such as educators, policymakers, or researchers through the National Opinion Research Center (NORC).

The dataset used for this paper was retrieved from The General Social Survey (GSS) Data Explorer website (citation). We retrieved survey data relating to work and job in the years of 1989, 1998, 2006, and 2016.

### 2.1 Source Data and Methodology

Majority of the GSS data was collected through face-to-face interviews with the target population of adults (18+) residing in the United States, but starting in 2002, Computer-assisted personal interviewing methods were introduced (3).

All the survey data used was in relation to job and work in the Work Orientation Module; the specific variable names extracted from the dataset being `intjob`, `hlpoths`, and `hlp soc`.

### 2.2 Data Cleaning

The open source statistical programming language R was used to clean and analyze the data, along with producing the graphs. The main packages that supported this process included `tidyverse`, `ggplot2`, `knitr`, `kableExtra`, `here`...

The cleaning process involved filtering the specific data variables used for our analysis from the downloaded GSS dataset, and renaming any variables with meaningful names. For example, rather than “`intjob`” being the column name for “importance of interesting work in a job”, we renamed it to `interesting_work`, as shown in Table #. Further, the numerical values representing the participants’ responses (1-5) were changed to the representative words/phrases (not important, very important, etc.).

Table 1: yus

Variable	NewName	Description	ExampleResponse
<code>intjob</code>	<code>interesting_work</code>	Importance of interesting work in a job	Very Important
<code>hlpoths</code>	<code>helping_others</code>	Importance of helping others in a job	Neither
<code>hlp soc</code>	<code>social_usefulness</code>	Importance of social usefulness in a job	Not Important

Variable	NewName	Description	ExampleResponse
----------	---------	-------------	-----------------

## 2.3 Data Terminology

The response choices for each question and their respective code in brackets are as follows: Inapplicable (-100), No Answer (-99), Do Not Know/Cannot Choose (-98), Very Important (1), Important (2), Neither (3), Not important (4), and Not Important At All (5). For our graphs, we did not include the Inapplicable, No Answer, and Do Not Know/Cannot choose responses to focus on the discernible participant responses.

## 2.4 Respondent Demographics

Figure 1 shows the number and percentage of male and female respondents for 1989, 1998, 2006, and 2016. Figure 2 shows the mean, median, mode, min, and max of respondent age for all four years of data collection.

``summarise()`` has grouped output by 'year'. You can override using the ``groups`` argument.

Table 2: Respondent Gender Demographics by Year

year	sex	count	percentage
1989	female	786	56.51
1989	male	605	43.49
1998	female	678	58.80
1998	male	475	41.20
2006	female	807	53.48
2006	male	702	46.52
2016	female	766	52.22
2016	male	701	47.78

Figure 1: ?(caption)

Table 3: Respondent Age Demographics by Year

Year	Mean	Median	Mode	Min	Max
1989	45	42	28	18	89
1998	45	42	33	18	89
2006	47	46	47	18	89
2016	49	49	58	18	89

Figure 2: ?(caption)

## 2.5 Graphs of Responses

Figure 3, Figure 4, and Figure 5, shows the responses to the prompt “On the following list there are various aspects of jobs. Please circle one number to show how important you personally consider it is in a job” where each graph represents one of the aspects. Respondents answered on a 1 to 5 likert scale where 1 represents “very important” and 5 represents “not important at all”.

### 2.5.1 Helps Others

In Figure 3, the proportion of respondents to the prompt “A job that allows someone to help other people?” is displayed. From the first year of data collection in 1989 to 2006, “Important” was the most selected response. In 2016, “Very Important” surpassed “Important” by 1%. In general, you can see an increase in “Very Important” respondents across the years while there is little change in the proportion of “Not Important” and “Not Important At All” responses. Further, there is a general decrease in “Neither” responses from 1989 - 2006 which is interrupted when there is a slight increase in 2016.

### 2.5.2 Interesting Work

Figure 4 shows the proportion of responses for the prompt “An interesting job?”. In 1989, “Important” responses was most chosen at around 50%. The following survey, in 1998, showed an increase in “Very Important” responses, where it had a similar proportion to “Important” responses 1989 and “Important” decreased to a proportion similar to “Very Important”.

### 2.5.3 Social Usefulness

Figure 5 displays the proportion of responses for the prompt “A job that is useful to society?”. There is a large increase in the proportion of “Very Important” responses from 1989 to 2016.

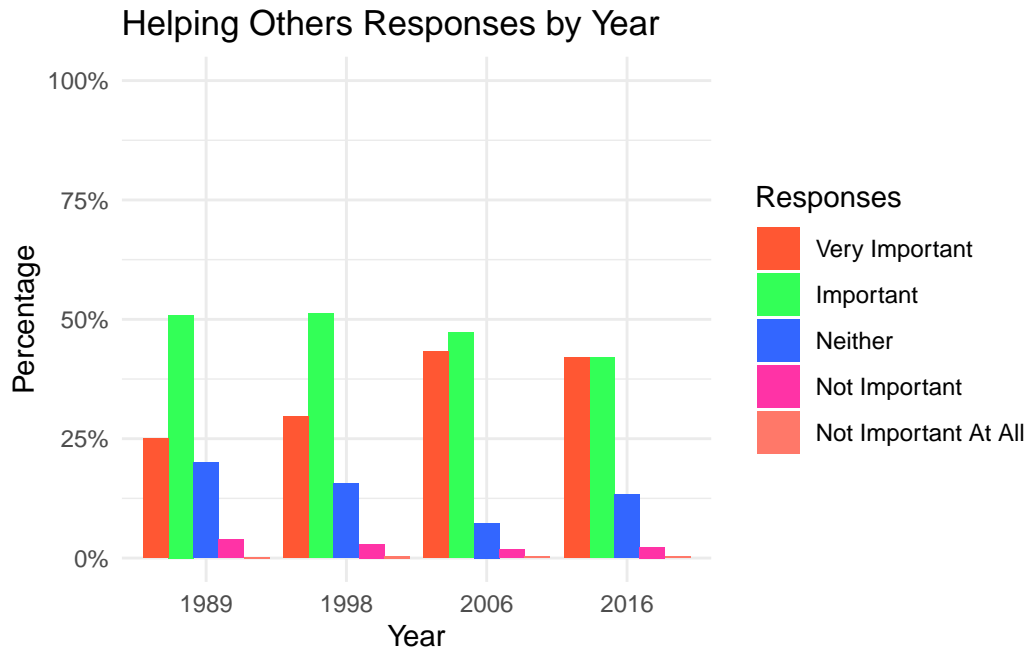


Figure 3: Q1 - “A job that allows someone to help other people?”

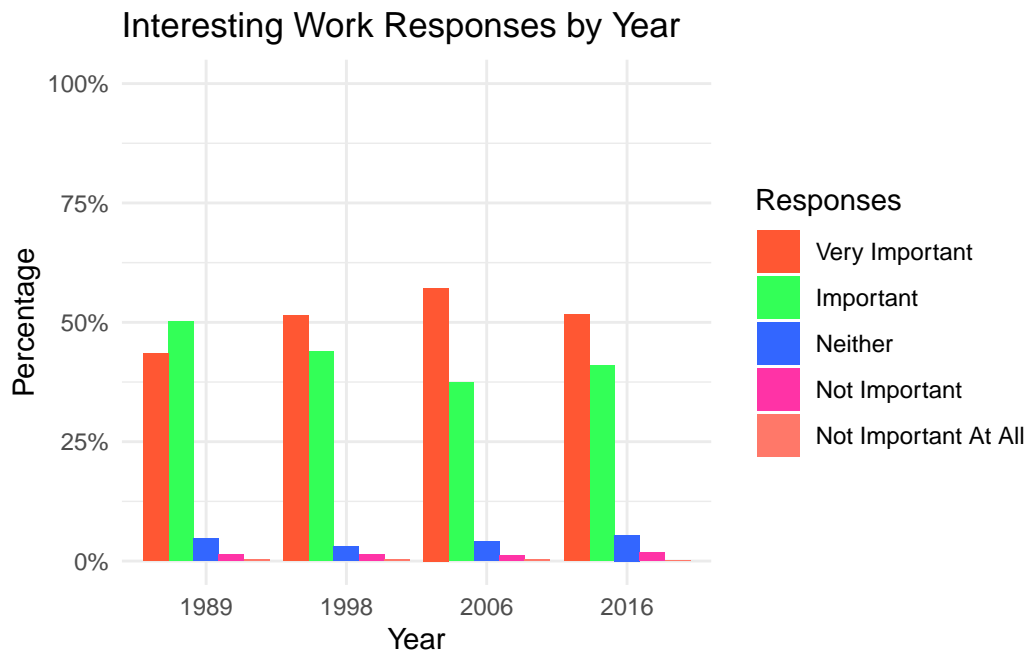


Figure 4: Q2 - “An interesting job?”

In contrast, there is a gradual decline for both “Important” and “Neither”. There is little change in “Not Important” and “Not Important At All”. Compared to the other figures, this graph has the most varying change in the “neutral” response.

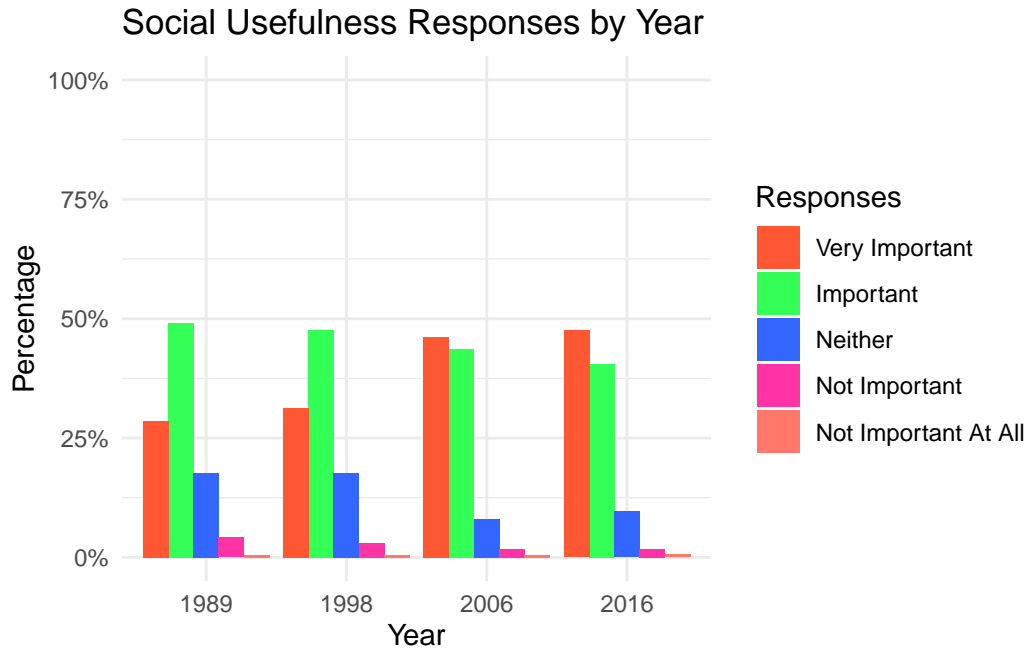


Figure 5: Q3 - “A job that is useful to society?”

### 3 Results

Table 2 summarizes the average of responses per year for each variable.

Table 4: Average of Responses by Year

Year	Helping Others	Interesting Work	Social Usefulness
1989	2.034	1.649	1.986
1998	1.932	1.550	1.939
2006	1.687	1.501	1.669
2016	1.767	1.577	1.675

``summarise()`` has grouped output by 'response'. You can override using the ``groups`` argument.

Table 5: Social Usefulness Reponse Proportions 2016

Response	Sex	Count	Percentage
very important	female	1228	57.09
very important	male	923	42.91
important	female	1389	55.90
important	male	1096	44.10
neither	female	339	47.68
neither	male	372	52.32
not important	female	67	46.53
not important	male	77	53.47
not important at all	female	14	48.28
not important at all	male	15	51.72

While all columns show a general increase in values, the Social Usefulness column is the only column that consistently conveys an increase.

?@fig-allresponse-helpingothers demonstrates the changing proportion of responses to how important is it for a job to help others.

```
#|label: fig-allresponse-helpingothers
#|echo: false
#|message: false
#|tbl-cap: Proportion of Responses to Importance of Job that Helps Other

# Prepare data
helping_others_data <- cleaned_GSS_data[c('year',
                                           'helping_others')]

helping_others_data <- helping_others_data |>
  mutate(
    helping_others = recode(helping_others,
                           '1' = 'Very Important',
                           '2' = 'Important',
                           '3' = 'Neither',
                           '4' = 'Not Important',
                           '5' = 'Not Important At All'),
  )

helping_others_data <-
  helping_others_data |>
```

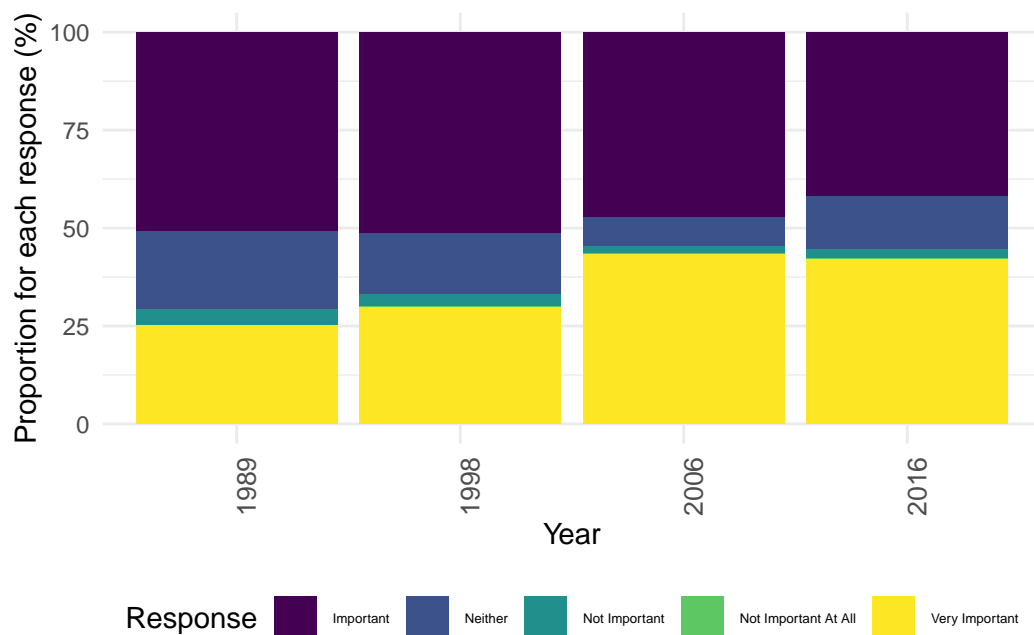
```

mutate(
  year = (as.character(year)),
)

helping_others_percentages <-
  helping_others_data |>
  count(year, helping_others) |>
  group_by(year) |>
  mutate(percentage = n / sum(n) * 100)

# Graph data
helping_others_percentages |>
  ggplot(mapping = aes(x = year, y = percentage, fill = helping_others)) +
  geom_col() +
  theme_minimal() +
  theme(axis.text.x = element_text(angle=90, size = 10)) +
  labs(x = "Year", y = "Proportion for each response (%)", fill = "Response") +
  theme(legend.position = "bottom") +
  theme(legend.text = element_text(size = 5)) +
  scale_fill_viridis_d(option = "viridis")

```





```

#|label: fig-allresponse-helpingothers
#|echo: false
#|message: false
#|tbl-cap: Proportion of Responses to Importance of Job that Helps Other

# Prepare data
helping_others_data <- cleaned_GSS_data[c('year',
                                           'helping_others')]

helping_others_data <- helping_others_data |>
  mutate(
    helping_others = recode(helping_others,
                           '1' = 'Very Important',
                           '2' = 'Important',
                           '3' = 'Neither',
                           '4' = 'Not Important',
                           '5' = 'Not Important At All'),
  )

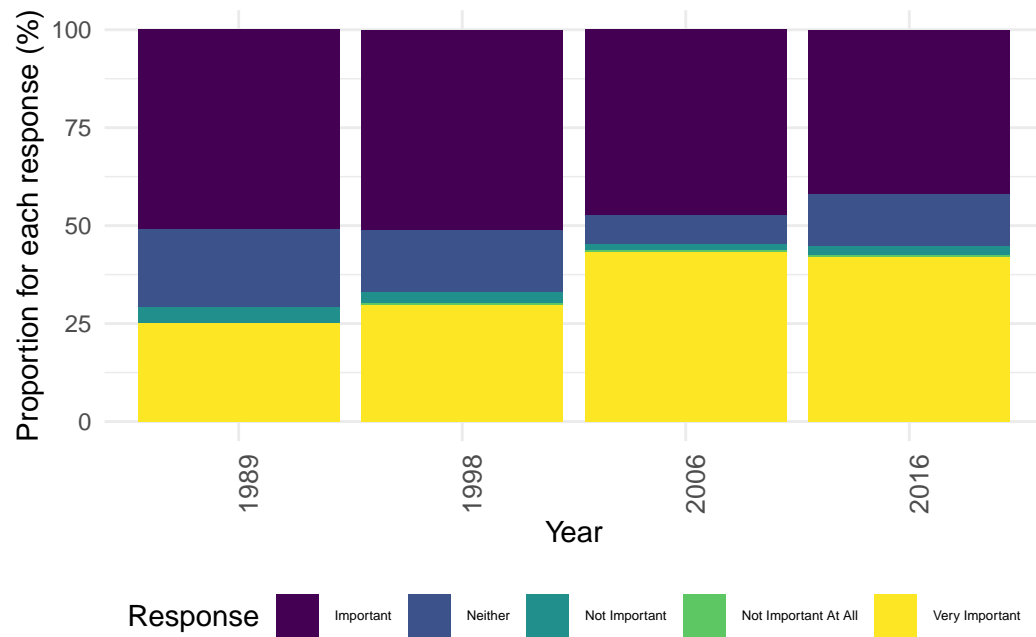
helping_others_data <-
  helping_others_data |>
  mutate(
    year = (as.character(year)),
  )

helping_others_percentages <-
  helping_others_data |>
  count(year, helping_others) |>
  group_by(year) |>
  mutate(percentage = n / sum(n) * 100)

# Graph data
helping_others_percentages |>
  ggplot(mapping = aes(x = year, y = percentage, fill = helping_others)) +
  geom_col() +
  theme_minimal() +
  theme(axis.text.x = element_text(angle=90, size = 10)) +
  labs(x = "Year", y = "Proportion for each response (%)", fill = "Response") +
  theme(legend.position = "bottom") +
  theme(legend.text = element_text(size = 5)) +

```

```
scale_fill_viridis_d(option = "viridis")
```



## 4 Discussion

### 4.1 Gender?

### 4.2 Culture?

## 5 Sources