Chicken Sandwich Preference by Gender, Age, and Region

Clayton Ramsey

December 2021

Contents

This is an investigation and analysis of survey data regarding the favorite chicken sandwich reported by individuals in the US.

Data description

The response variable to this survey is in the variable, "Question #1 Raw Response." The variable called "Synonym Group" is basically the same thing. The second table below shows the only two rows in which those variables are different. Classifying them using regex without case sensitivity, they would be identical.

Question #1 Raw Response	Total
Chick-Fil-A	768
Popeyes	323
Wendy's	180
Kentucky Fried Chicken (KFC)	154
McDonald's	118
Burger King	97
Me	14
None	7
me	5
Mom	3
I'm a vegetarian	2
My wife	2
Myself	2
none	2
Sheetz	2

Question #1 Raw Response	Question #1 Synonym Group
me	Me
Me!	Me

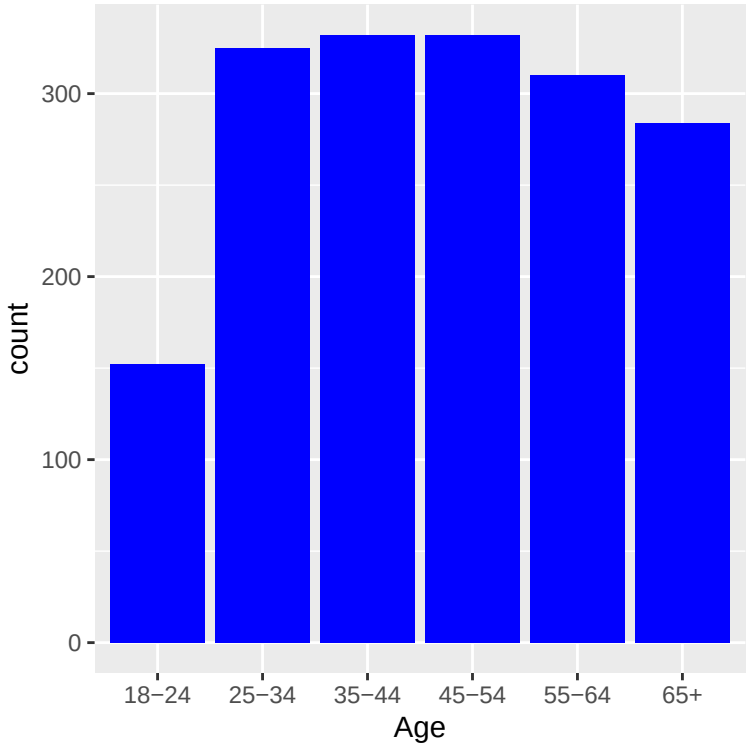
The responses fall into several categories:

- Six major chains, like Chick-Fil-A and Burger King.
- Smaller restaurants, most likely regional or local places, like Pollo Campero or Lees, and major chains not known for their chicken, like Hardee's
- Responses indicating the respondent prefers home-cooked chicken. ("Me.", "My Grandma")
- Responses indicating the respondent does not or cannot eat chicken. ("Stop eating chickens.", "allergic to chicken", "I'm a vegetarian")
- Responses indicating the respondent doesn't eat fast food, or is indifferent to the survey. ("Don't eat trash", "Who cares")
- Variations on "none"

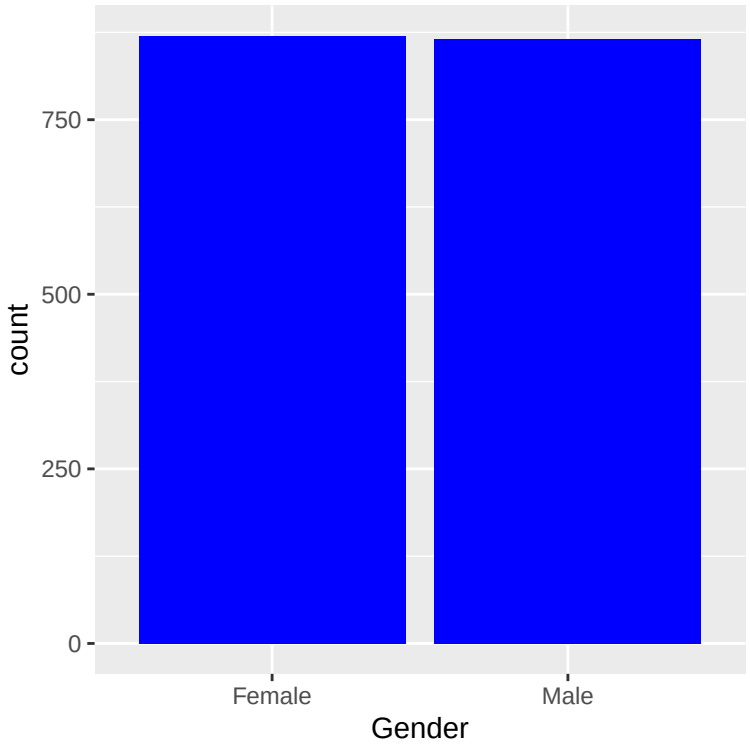
This data could use a transformation of the response variable. The top six responses by frequency are the major chains, so these should be left alone. I will also create new categories "Home Cooked", "None", and "Other Restaurant".

Response	Total
Chick-Fil-A	768
Popeyes	323
Wendy's	180
Kentucky Fried Chicken (KFC)	154
McDonald's	118
Burger King	97
Home Cooked	40
Other	28
None	27

The "Age" variable is a factor of age ranges.



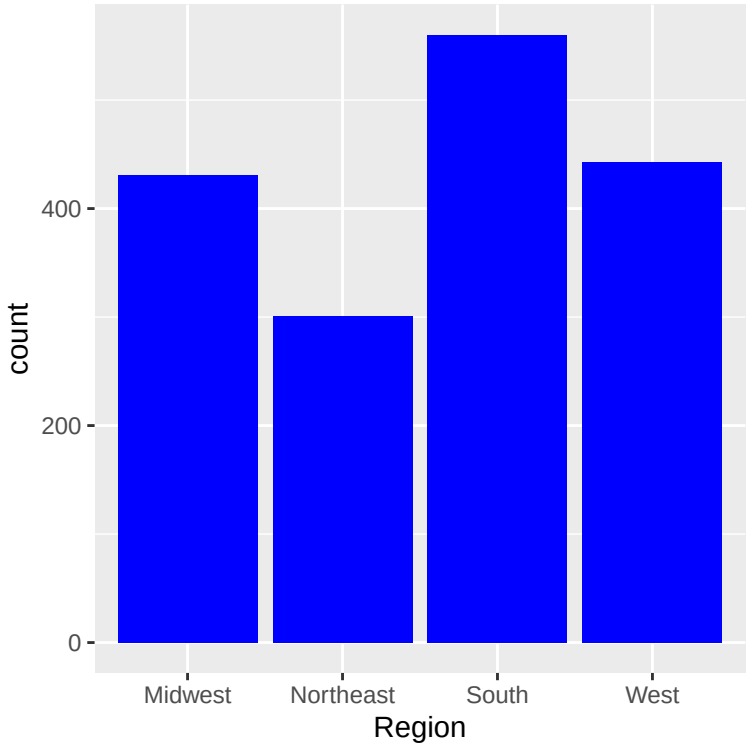
The “Gender” Variable shows that the responses were balanced.



The “Geography” variable includes localities within individual states, meaning it has a large number of unique values. Here are the first 15 entries.

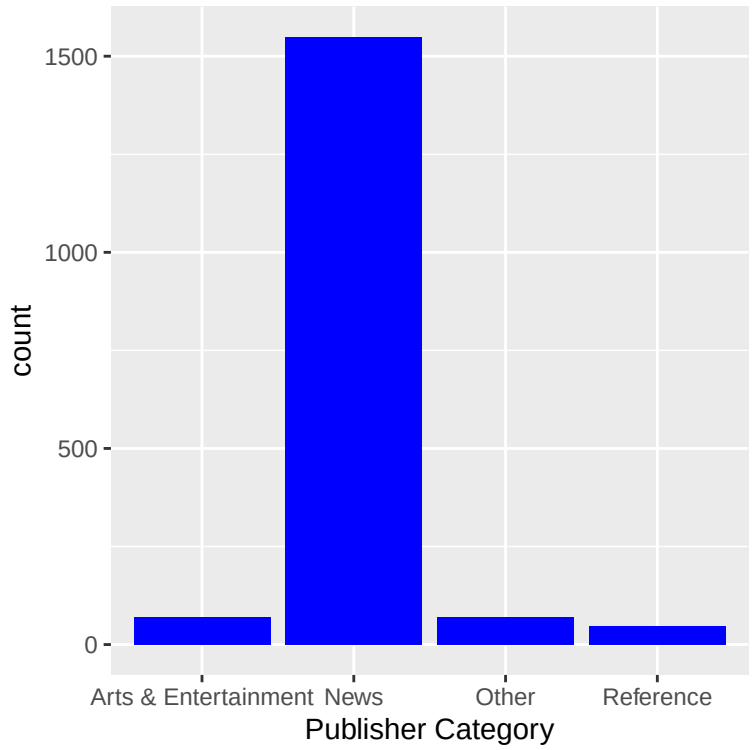
Geography	Total
US-MIDWEST-IA	27
US-MIDWEST-IA-Omaha	5
US-MIDWEST-IL	26
US-MIDWEST-IL-Chicago	15
US-MIDWEST-IN	28
US-MIDWEST-IN-Indianapolis	10
US-MIDWEST-KS	32
US-MIDWEST-MI	30
US-MIDWEST-MN	25
US-MIDWEST-MN-Minneapolis	7
US-MIDWEST-MO	35
US-MIDWEST-MO-KCMO	11
US-MIDWEST-ND	33
US-MIDWEST-NE	26
US-MIDWEST-NE-Lincoln	9

Fortunately, the strings in this column contain region tags so I can create a new “region” variable using regex.

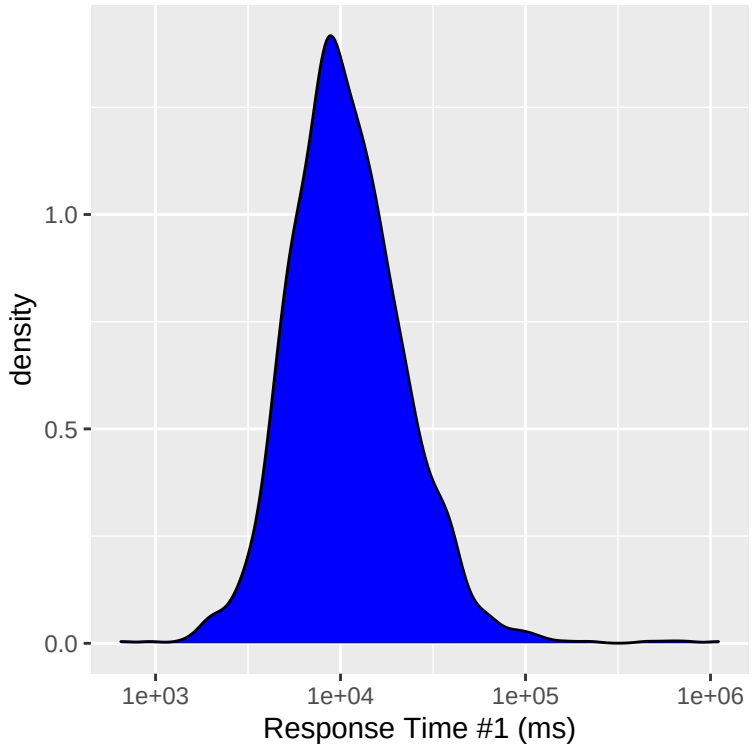


Region	Total
Midwest	431
Northeast	301
South	560
West	443

The "Publisher Category" has to do with how the respondent encountered the survey.



The "response time" variable measures the time to survey completion in ms.



Descriptor	Value
Min	649.0
Q1	6810.0
Median	10266.0
Q3	16608.5
Max	1105000.0

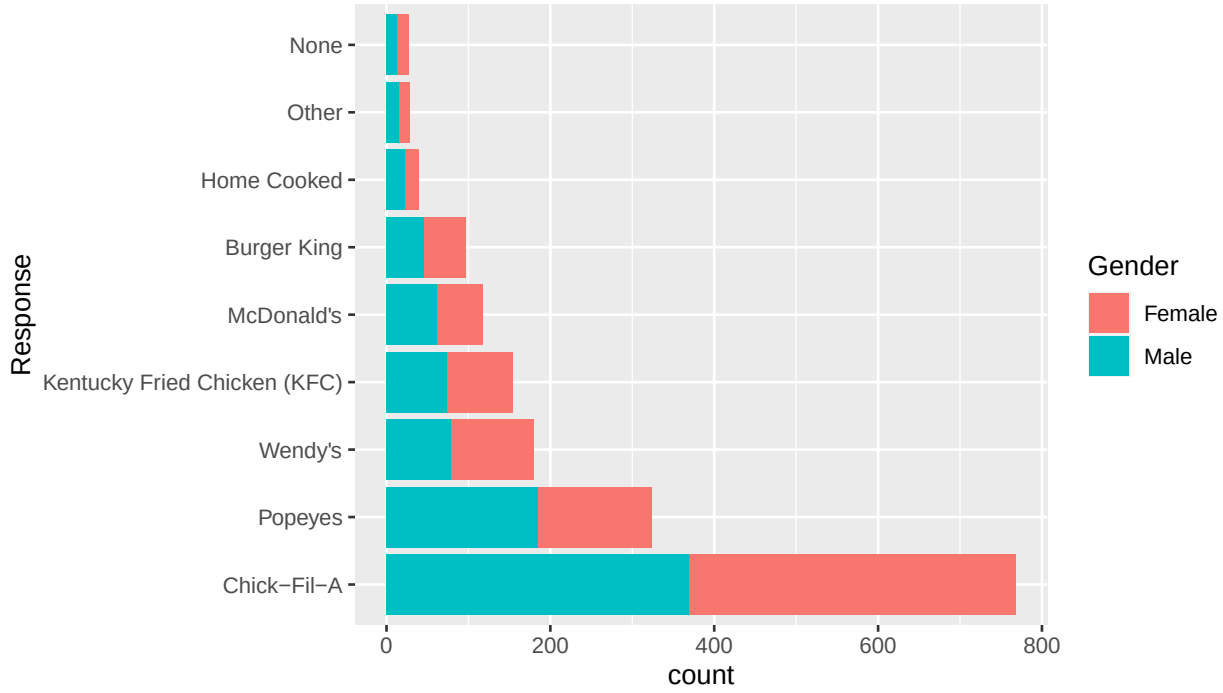
The mean value is 1.5726255×10^4.

The final variable of interest is "state", which gives the number of respondents from each one.

state	Total
AK	32
AL	32
AR	31
AZ	36
CA	41
CO	36
CT	30
DC	31
DE	31
FL	32

Analysis

The first question I wanted to consider was whether different chicken sandwiches were preferred significantly by a certain gender. Here we see the responses colored by gender.



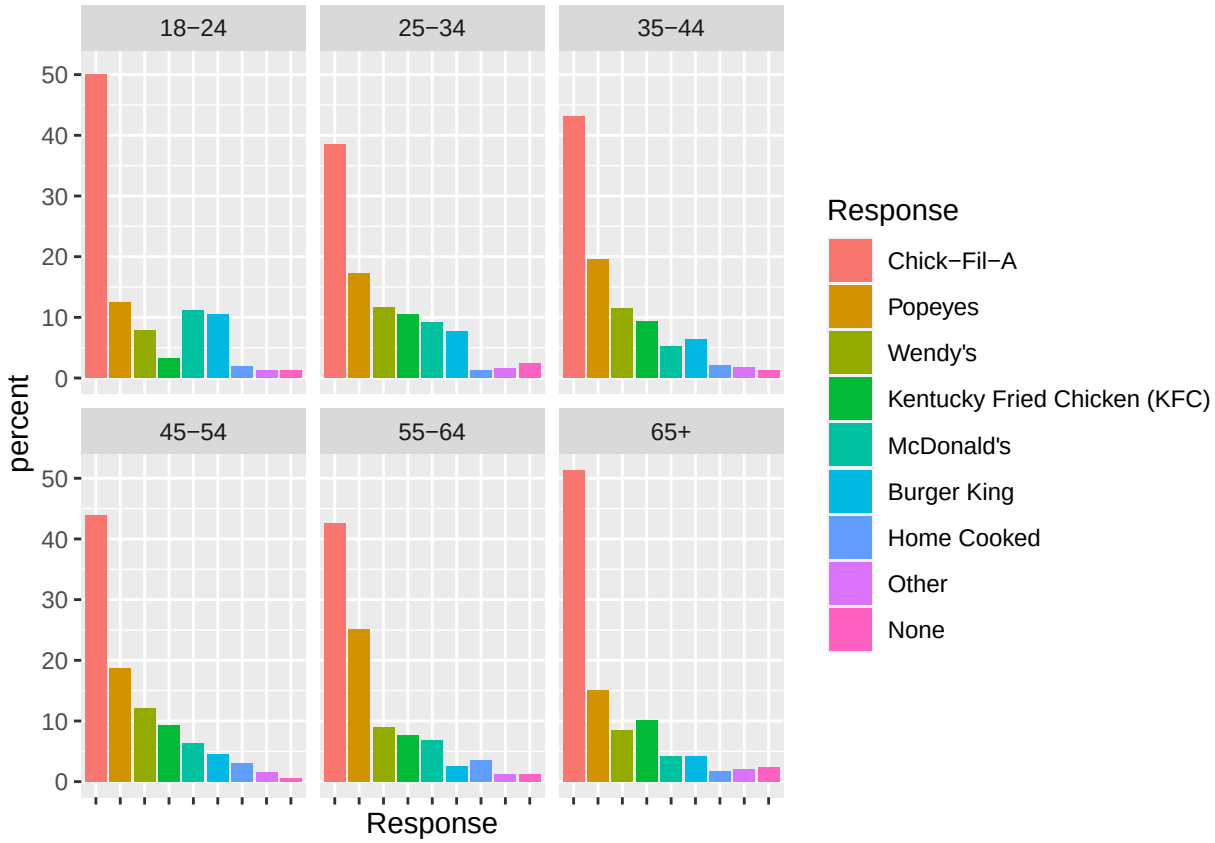
These all look like the gender balance is the same, but I conducted Binomal Tests just to be sure, using a null hypothesis that the genders are equally represented, P(M) = P(F) = 0.5. For these tests, I will form confidence intervals on the Female proportion.

Response	Lower	Upper	Gender Balance
Chick-Fil-A	0.4835279	0.5553841	Yes
Popeyes	0.3726358	0.4832000	No
Wendy's	0.4853400	0.6348360	Yes
Kentucky Fried Chicken (KFC)	0.4376195	0.6005776	Yes
McDonald's	0.3819136	0.5685454	Yes
Burger King	0.4218204	0.6281102	Yes
Home Cooked	0.2925884	0.6150932	Yes
Other	0.2751086	0.6613009	Yes
None	0.3194965	0.7133275	Yes

The only category with evidence of a gender-imbalanced group of respondents was Popeyes, showing a slight male preference.

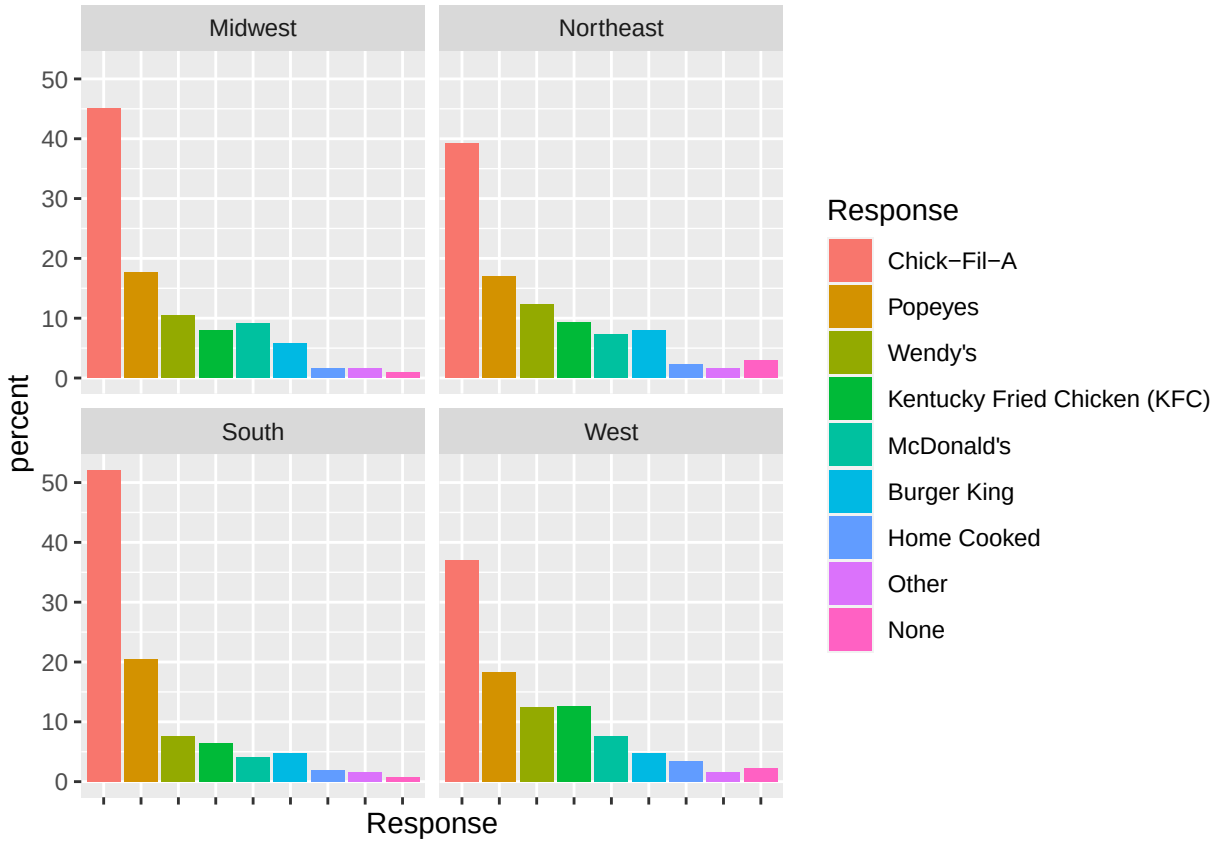
Next, I wanted to look at the preferred sandwich in the different age bands. The plot below shows the responses organized by age group and graphed as the percent of total responses. There are several insights to be gained from this data:

- Chick-Fil-A is the most popular for every age group.
- Popeye's comes in a clear second in every category except the 18-24 year olds, for whom McDonald's and Burger King are competitive.
- 45-54 year olds show preferences ranked the same way as the dataset as a whole. (The legend is listed in frequency order.)
- The categories "Other" and "None" look as though they have the highest relative popularity in the 25-34 and 65+ categories.



Finally, I wanted to consider the popularity of different chicken sandwiches according to region. The results are graphed below, also by percent of total responses. Again, there are insights to be found:

- The relative popularity of Chick-Fil-A is highest in the South and lowest in the West. This stands to reason somewhat, since these restaurants are less common in the western US.
- The Midwest is the only region where McDonald's is more popular than KFC.
- Looking at the three free-response categories, we see that:
  - The "None" and "Home Cooked" categories are more popular that "Other" in the Northeast and West.
  - The "Other" category shows a consistent response of about 2% in all four regions.
  - In the Midwest and South, "Other" is comparable to "Home Cooked".
  - "None" is the most popular of the three only in the Northeast.


Midwest
Northeast
South
West
percent
50
40
30
20
10
0
Response
Response
Chick−Fil−A
Popeyes
Wendy's
Kentucky Fried Chicken (KFC)
McDonald's
Burger King
Home Cooked
Other
None