

# Thanksgiving Celebration Habits

Based on data from a 2015 US survey



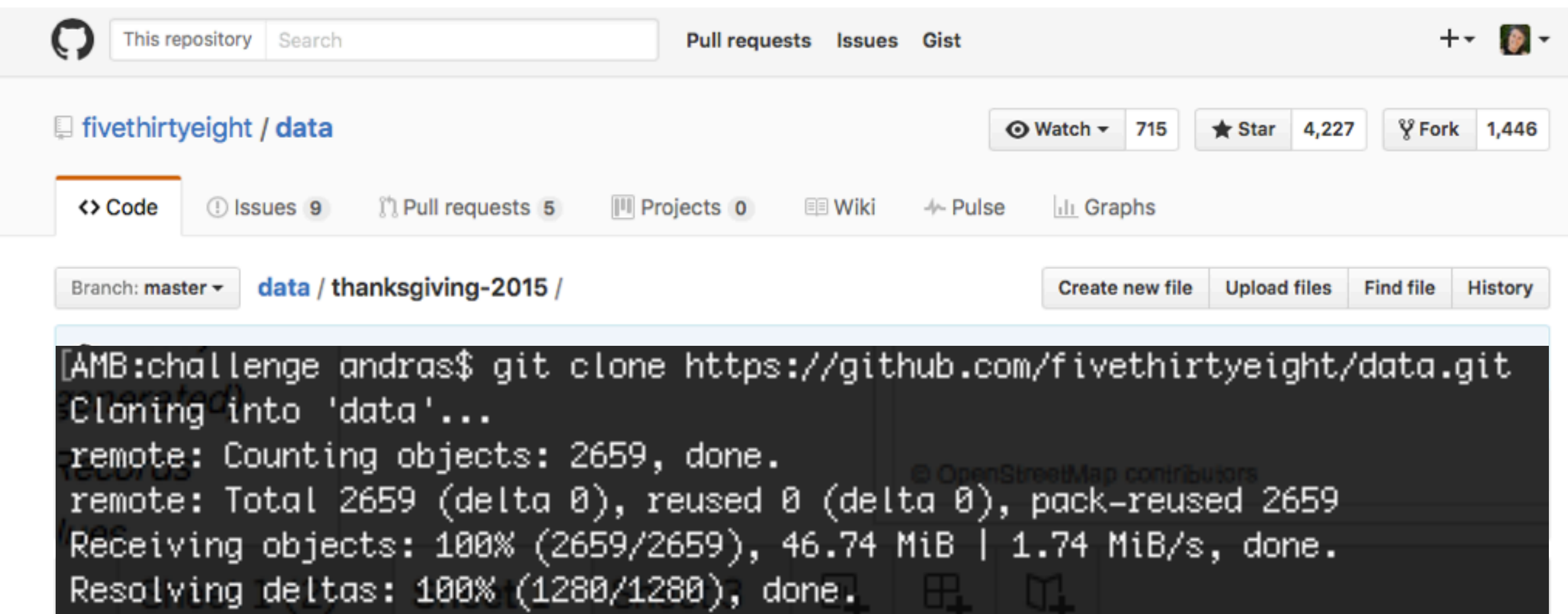
# Introducing the dataset

- Thanksgiving is a national holiday celebrated in Canada and US
- Using a SurveyMonkey poll, 1,058 respondents answered questions about their Thanksgiving in 2015
- The dataset contains basic personal information (age, sex, income, region) combined with eating/socialising habits (consumed food, going out, TV) on Thanksgiving day

What we did...

# 1. Getting the data

- using everyone's favourite new tool, git



The image shows a screenshot of a GitHub repository page for 'fivethirtyeight / data'. The repository has 715 watches, 4,227 stars, and 1,446 forks. The 'Code' tab is selected, showing the 'master' branch. Below the repository information, there is a terminal window displaying the output of a 'git clone' command. The terminal text is as follows:

```
[AMB:challenge andras$ git clone https://github.com/fivethirtyeight/data.git
Cloning into 'data'...
remote: Counting objects: 2659, done.
remote: Total 2659 (delta 0), reused 0 (delta 0), pack-reused 2659
Receiving objects: 100% (2659/2659), 46.74 MiB | 1.74 MiB/s, done.
Resolving deltas: 100% (1280/1280), done.
```

## 2. Cleaning the data

- data was fairly clean and well organised
- exploratory data analysis was conducted using Tableau
- during this process we filtered the useful columns for our analysis

### Importing csv file into list

```
with open('w01-5.1-thanksgiving-2015-poll-data.csv', 'rU') as csvfile:
    reader = csv.reader(csvfile)
    rows = list(reader)

#for r in rows[0]:
#    print r[:50]

print 'Initial number of columns:', len(rows[0])
print 'Total number of respondents:', len(rows)-1
print '=====

for r in rows[0][:2]+rows[0][-4:]:
    print r
```

```
Initial number of columns: 65
Total number of respondents: 1058
=====
RespondentID
Do you celebrate Thanksgiving?
Age
What is your gender?
How much total combined money did all members of your HOUSEHOLD earn last
US Region
```

# 3. Analysing the data

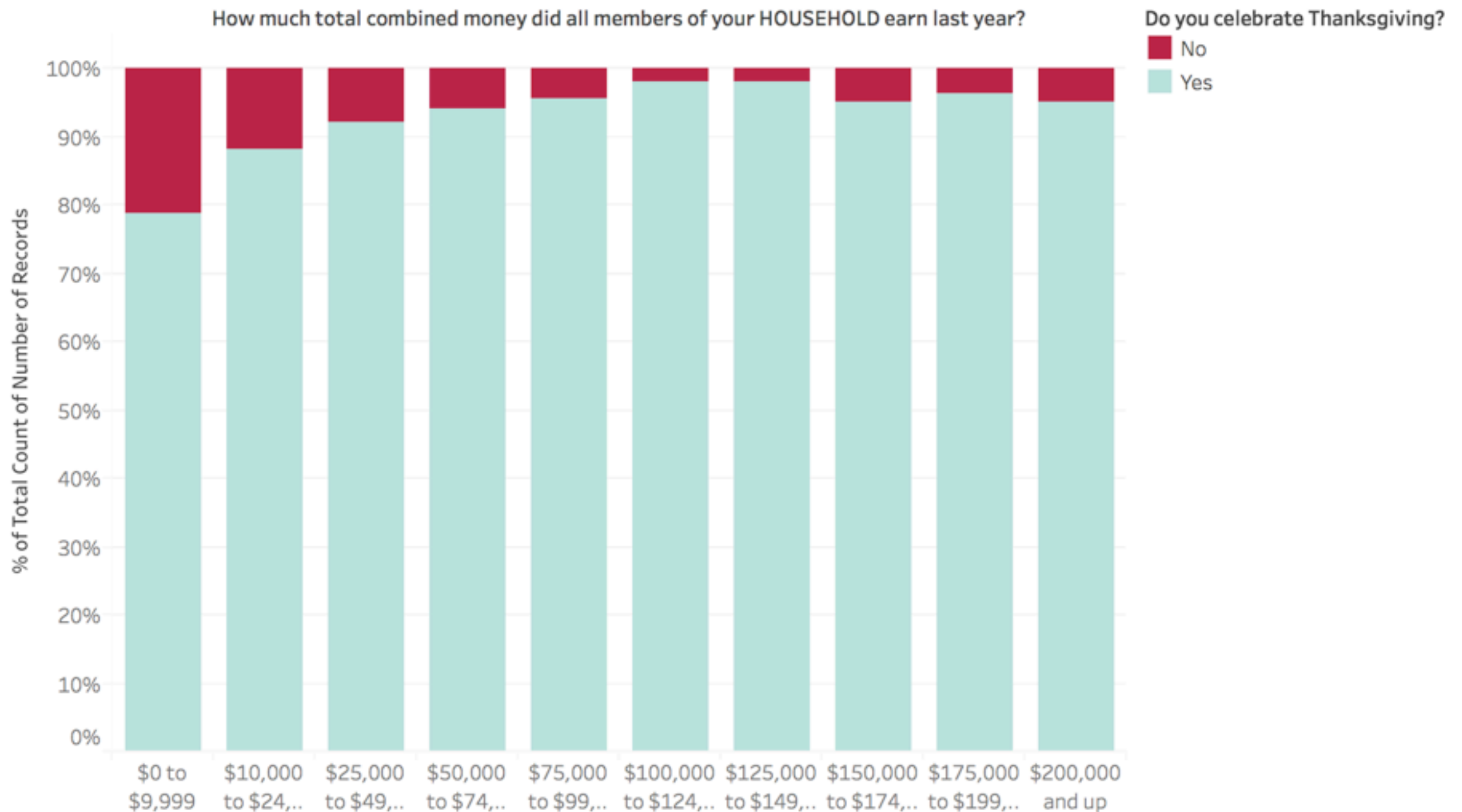
- using



# 4.1 Visualising the data

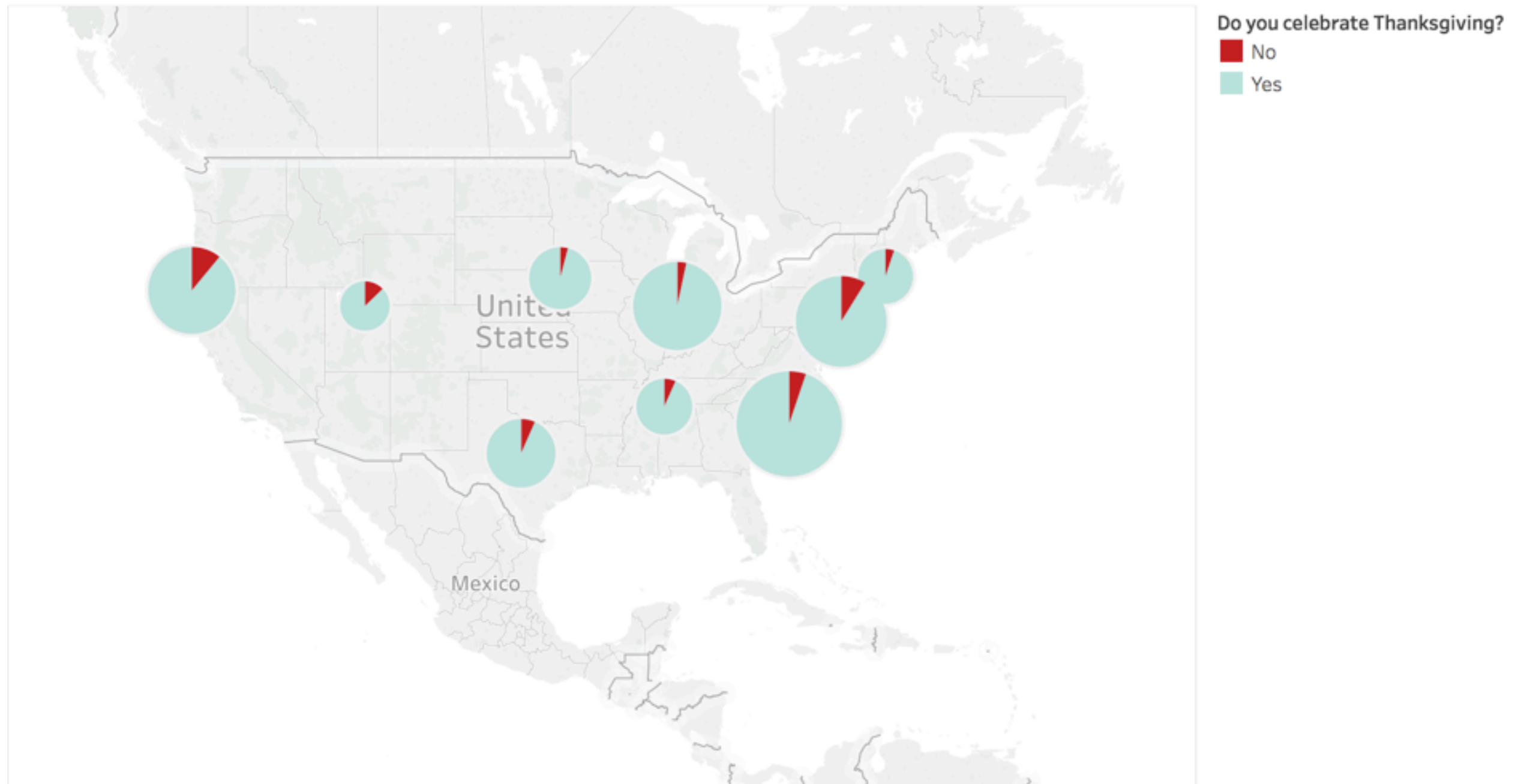
## The Thanksgiving - income dimension

Proportion of people celebrating Thanksgiving based on their annual income



## 4.2 The Thanksgiving - region dimension

Celebrating Thanksgiving in different areas



## 4.3 Celebrating Thanksgiving

Proportion of people participating to Thanksgiving according to gender and age.



Proportion of people that pray before eating Thanksgiving dinner according to area





## 4.4 Where people spend Thanksgiving

Proportion of people travelling for thanksgiving according to Age

Age

18 - 29

30 - 44

45 - 59

60+



How far will you travel for Thanksgiving?

- Thanksgiving is happening at my home--I won't travel at all
- Thanksgiving is local--it will take place in the town I live in
- Thanksgiving is out of town and far away--I have to drive several hours or fly
- Thanksgiving is out of town but not too far--it's a drive of a few hours or less

Not long to go...



*Thanks!*