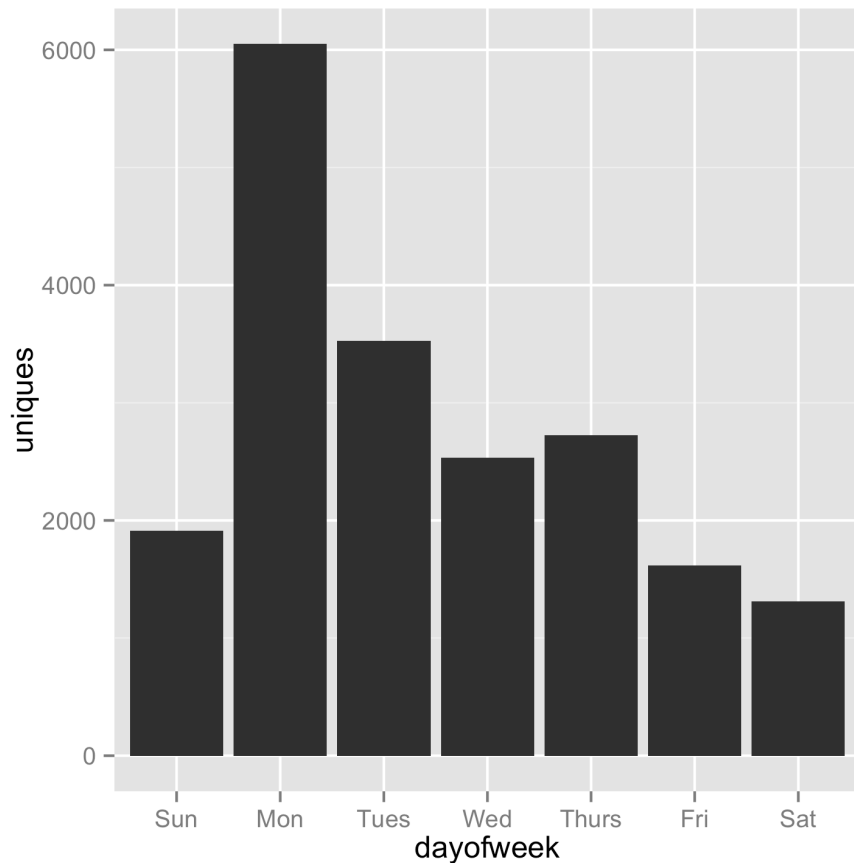


# How?

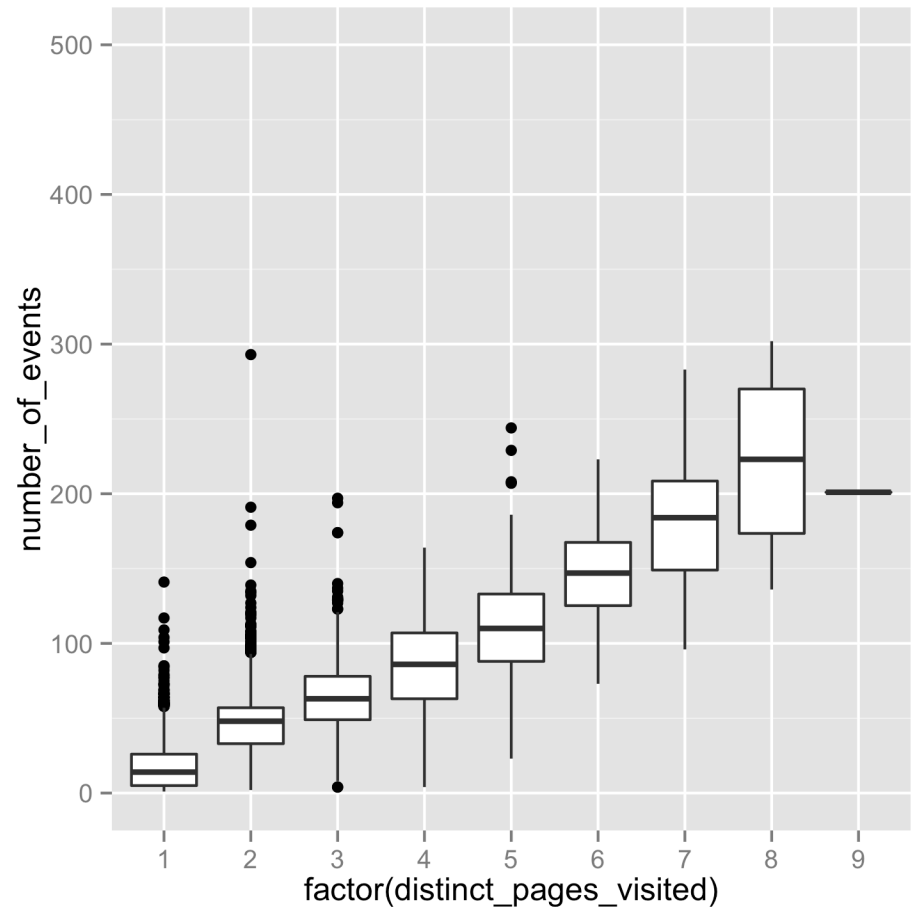
- We assume our data includes information about URLs with high bounce rates
- Most visits are from USA. We have neglected other countries in our analysis.
- All analysis and data manipulation performed in R. List of most used R libraries:
  - ggplot2, dplyr, tidyr, ggmap, arules, etc.

# Breadth and depth of user engagement

Unique visits per day

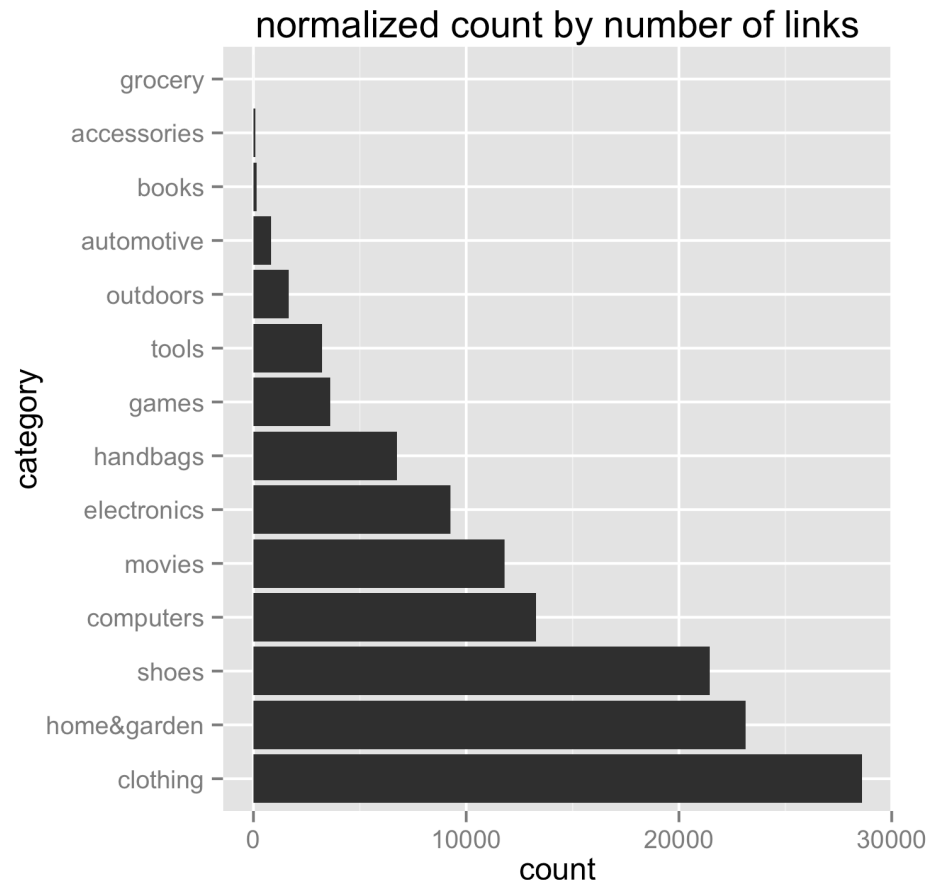


Engagement depth increases with the breadth



# Top three categories with highest bounce rates are Clothing, Home&Garden and Shoes

Clothes (6), handbags (8) and shoes (6) have multiple links in the data.  
Plot below shoes normalized count of different categories.

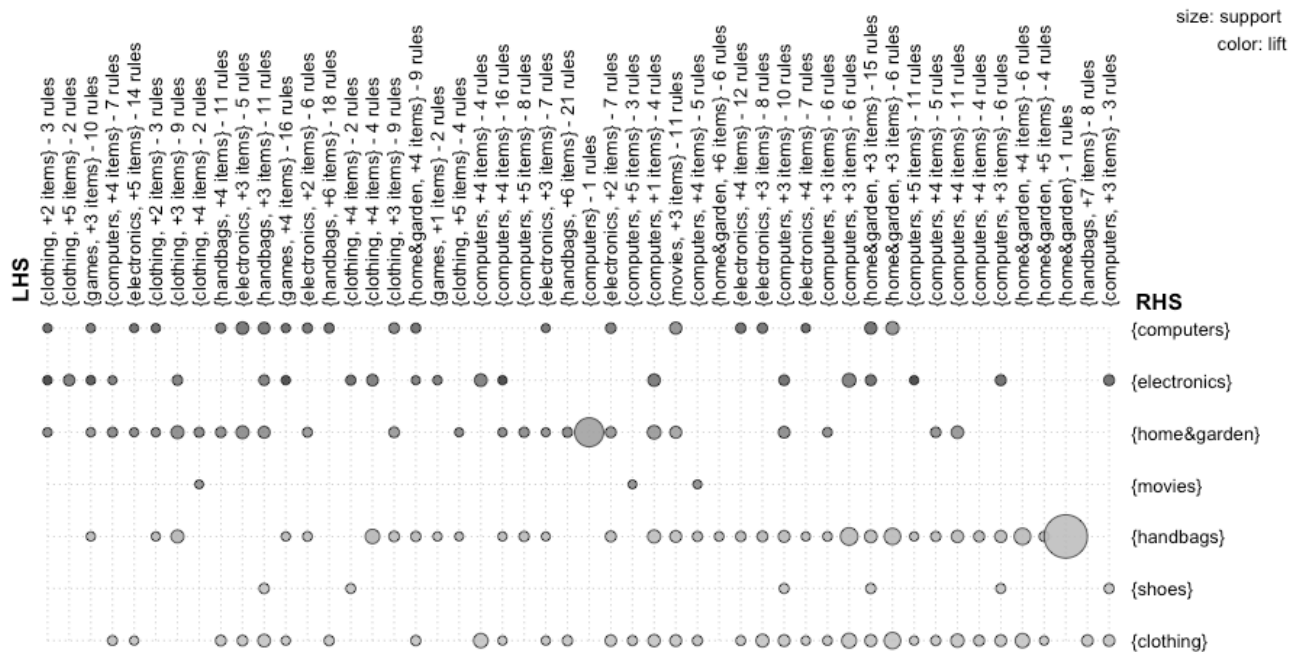


# Item frequently purchased together

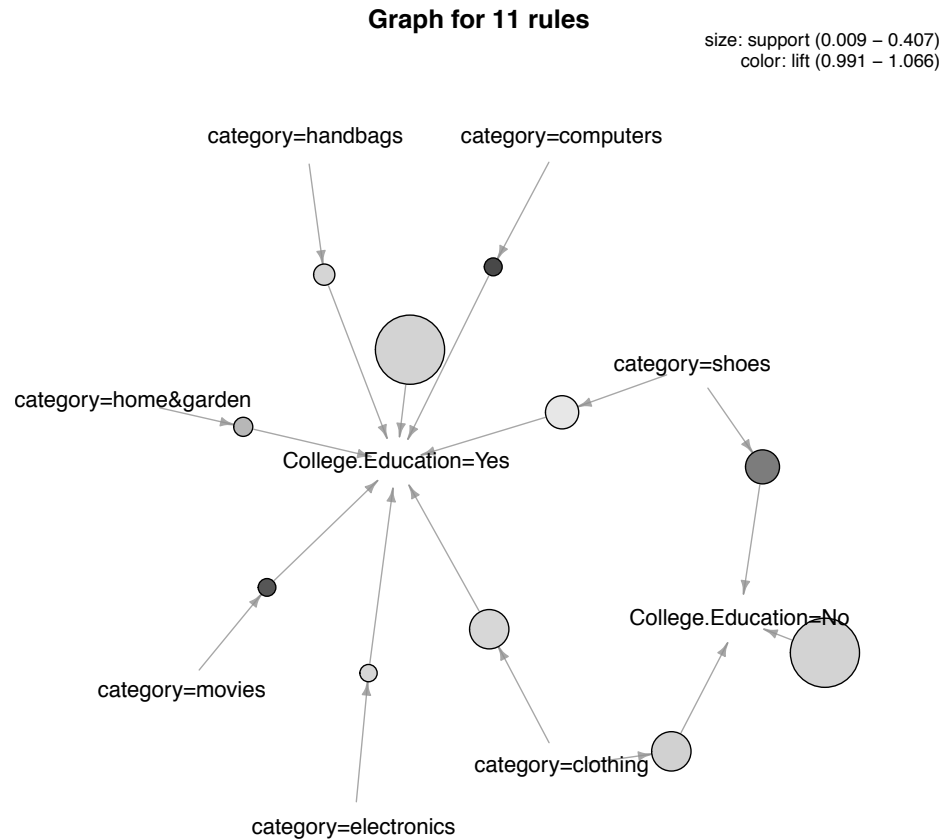
most frequent items in order of relative importance:

clothing    handbags    shoes  
3            2            2            1

Grouped matrix for 369 rules

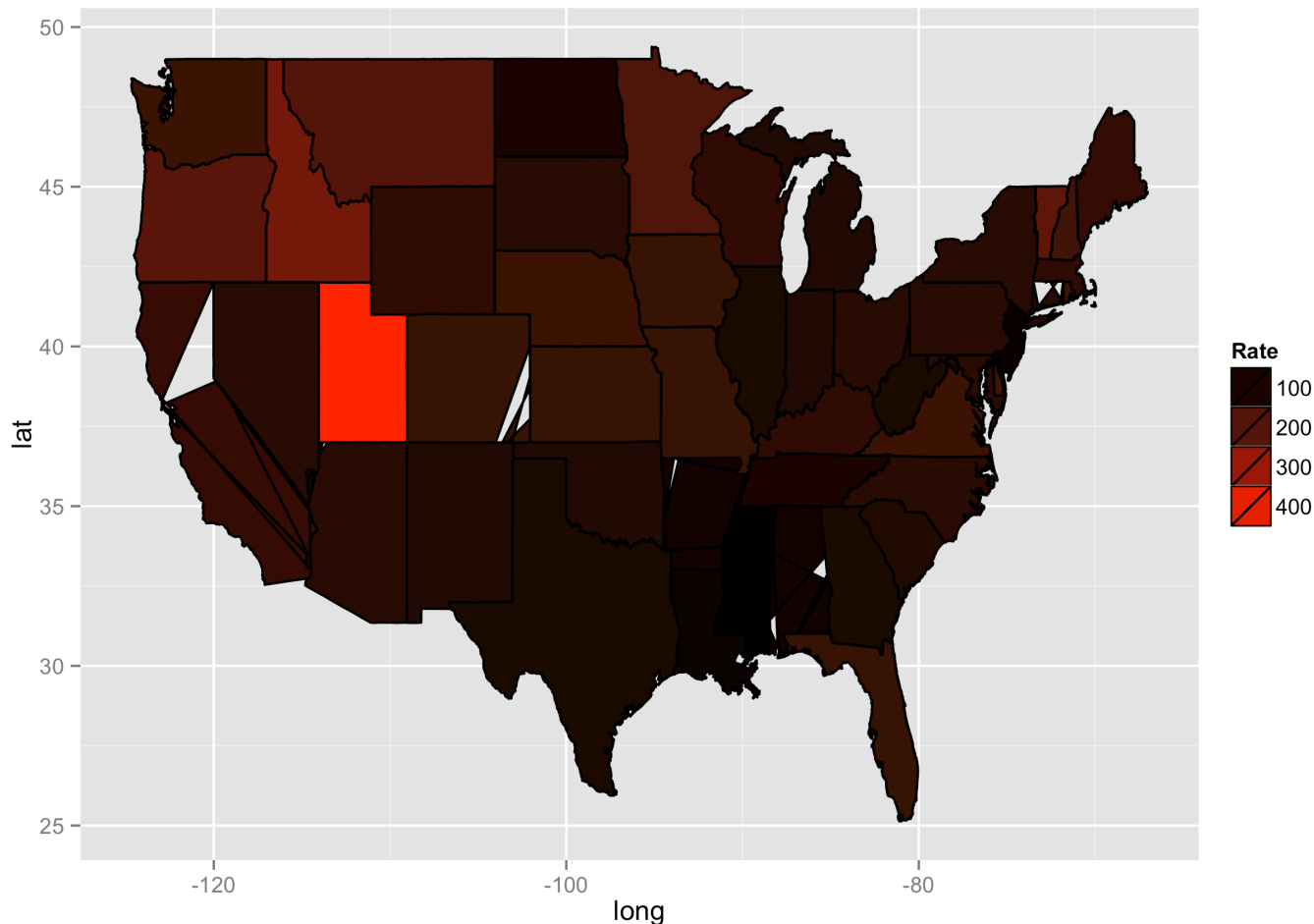


# People who have college education spend on more categories

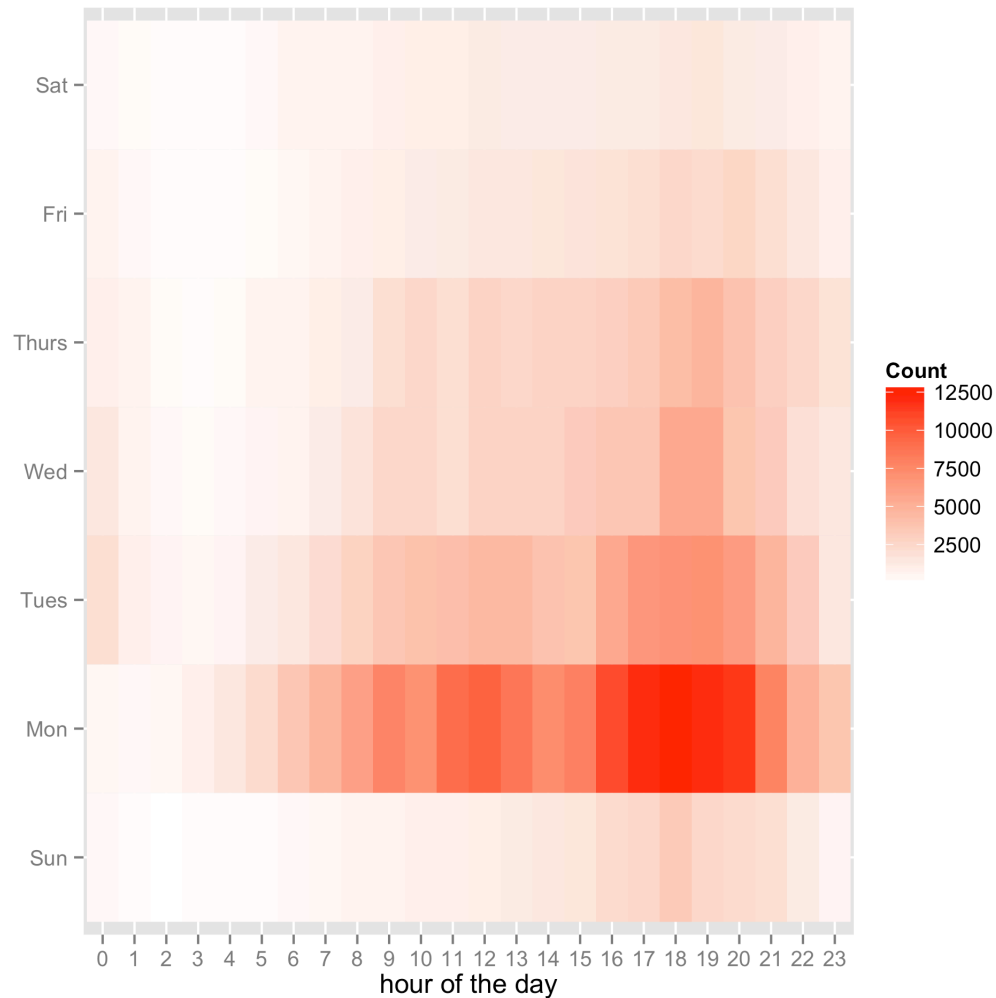


# Top 3 states with highest visit rate are Utah, Idaho and Oregon

Visit rate is Visits per 100,000 population of the state. In raw data California, Florida, etc. were top states therefore data was normalized by the state population.



# Peak day and time of user activity



# User activity broken down by categories and time and day

