

Amazon Mechanical Turk

Crowdsourcing and Human Computation
Lecture 2

Instructor: Chris Callison-Burch
TA: Ellie Pavlick

Website: crowdsourcing-class.org

Everyone's first questions about MTurk

- What type of tasks can be completed on MTurk?
- Who are the workers?
- How much money do they make?
- How much does it cost to get results?
- How fast can you get results back?
- How big is the MTurk marketplace?

Discussion of HW1

MTurk UX for Workers

All HITs | HITs Available To You | HITs Assigned To You

Find HIT

- HIT Creation Date (oldest first)
- HIT Creation Date (newest first)
- HITs Available (fewest first)
- HITs Available (most first)
- Reward Amount (least first)
- Reward Amount (most first)**
- Expiration Date (soonest first)
- Expiration Date (latest first)
- Title (A-Z)
- Title (Z-A)
- Time Allotted (least first)
- Time Allotted (most first)

GO!

Show all details | Hide all details

I understand, would be useful to know how to clean up aud

HIT Expiration Date: Sep 4, 2013 (1 day 21 hours) **Reward:**

Time Allotted: 24 hours **HITs Available:**

Transcription of approximately 15 minutes of audio

Requester: [Amelia Jones](#)

HIT Expiration Date: Sep 5, 2013 (2 days 23 hours) **Reward:**

Time Allotted: 8 hours **HITs Available:**

BY INVITATION ONLY: Answer a few questions in a brief survey

Requester: [Qualtrics Survey](#)

HIT Expiration Date: Sep 5, 2013 (2 days 21 hours) **Reward:**

Time Allotted: 30 minutes **HITs Available:**

Take a geo-tagged photo of a landmark in North Dakota

Requester: [Crowdsourcing at Thomson Reuters](#)

HIT Expiration Date: Sep 5, 2013 (2 days 16 hours) **Reward:**

Time Allotted: 3 days **HITs Available:**

HITs

containing []

that pay at least \$ **Dashboard - Chris Callison-Burch** (If you're not Chris Callison-Burch, [click here.](#))

Your Worker ID: A23KO2TP7I4KK2

Total Earnings ([What's this?](#))

Rewards You Have Earned	Value
Approved HITs	\$44.29
Bonuses	\$0.00
Total Earnings Show earnings details ▾	\$44.29

Your HIT Status ([What's this?](#))

Date	Submitted	Approved	Rejected	Pending	Earnings
Today	2	1	0	1	\$0.05
Sep 2, 2013	8	8	0	0	\$0.24
View more...					

HIT Totals ([What's this?](#))

HITs You Have Accepted	Value	Rate	HITs You Have Submitted	Value	Rate
HITs Accepted	1303	—	HITs Submitted	1095	—
... Submitted	1095	84.0%	... Approved	1081	98.8%
... Returned	119	9.1%	... Rejected	13	1.2%
... Abandoned	89	6.8%	... Pending	1	—

HITs

containing

that pay at least \$ 0.00

► require Master Qualificat

Status Detail - Chris Callison-Burch (If you're not Chris Callison-Burch, click here.)

Show: All HITs

HITs You Worked On For September 02, 2013 ([What's this?](#))

What sucks about the MTurk worker interface?

- No way to search by expected hourly rate
- No way to estimate the difficulty or length of time that it will take to do a task
- No way to know how reputable a requester is in advance

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[Results](#)[Workers](#)[Qualification Types](#)[Manage HITs](#)

Manage Batches

Click on the name of the batch to see more details

▼ Batches in progress (1)

Compression HIT - grammar/meaning 10		Results	Cancel th...
Created:	July 04, 2013	Assignments Completed:	2,468 / 2,468
Time Elapsed:	1 day	Estimated Completion Time:	COMPLETE
Average Time per Assignment:	3 minutes 40 seconds	Effective Hourly Rate:	\$4.091
Batch Progress:	<div style="width: 100%;"><div style="width: 100%;">100% submitted</div></div>	<div style="width: 100%;"><div style="width: 100%;">100% published</div></div>	

▼ Batches ready for review (143)

[« Previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) ...

Word Alignment - Trusted Workers - Dev 0.7		Results	Delete
Created:	July 02, 2013	Assignments Completed:	1,995 / 2,000
Time Elapsed:	15 days	Estimated Completion Time:	September 03, 2013 4:41 PM (TODAY)
Average Time per Assignment:	1 minute 58 seconds	Effective Hourly Rate:	\$4.576

HIT - grammar/meaning 10

this batch, make changes, or get results.

ces and decide how grammatical they are

[Cancel this batch](#)

100% submitted

100% published

2,468 / 2,468

July 04, 2013 1:03 PM PDT

Average Time per Assignment: 3 minutes 40 seconds

Estimated Completion Time: COMPLETE

Average Hourly Rate: \$4.091

HIT - grammar/meaning

Read a group of sentences and decide how grammatical they are
English, writing, editing, language, research, language, paraphrasing,
paraphrases, summarization, summaries, text, comprehension, reading
HIT Approval Rate (%) for all Requesters' HITs greater than or equal to 85
(Required for preview)

Number of HITs Approved greater than 100 (Required for preview)

Location is UNITED STATES (Required for preview)

er HIT: 2

\$0.250

compression-evaluation.dev.csv

Results

Assignments pending review: 0

Assignments approved: 2,070

Assignments rejected: 398

Resu

Cost Summary

Estimated Total Reward: \$500.000

Estimated Fees to Mechanical Turk: \$50.000 (fee details)

Estimated Total Cost: \$550.000

These costs are only an estimate until all of the assignments have been submitted and reviewed.

[Approve](#)[Reject](#)

<u>Worker ID</u>	<u>Work Time In Seconds</u>	Lifetime Approval Rate	Input.System Names 0	City	Region
A1H11P4QSM08Y1	145	100% (59/59)	pos_control neg_control /export/common/SCALE13/...		
A3O7D2TTK99UK5	3	12% (119/1007)	pos_control neg_control /export/common/SCALE13/...		
ASB13NJKTGCYB	122	100% (44/44)	pos_control neg_control /export/common/SCALE13/...		
A1BM57TUNAQIXM	478	100% (3/3)	pos_control neg_control /export/common/SCALE13/...	Los Angeles	California
A3445FT062DGIX	231	100% (78/78)	pos_control neg_control /export/common/SCALE13/...		
AYSX5GNO640MZ	141	13% (5/39)	pos_control neg_control /export/common/SCALE13/...		
A21U64TBRENETA	158	100% (17/17)	pos_control neg_control /export/common/SCALE13/...	Orlando	Florida
A23AFAXSEYH61S	241	100% (177/177)	pos_control neg_control /export/common/SCALE13/...	Los Angeles	California
A1GYXHW4YCOJ0H	154	100% (77/77)	pos_control neg_control /export/common/SCALE13/...		
A3O7D2TTK99UK5	21	12% (119/1007)	pos_control neg_control /export/common/SCALE13/...		
AUGRDUEDEUXUS	184	100% (12/12)	pos_control neg_control /export/common/SCALE13/...		
AYSX5GNO640MZ	182	13% (5/39)	pos_control neg_control /export/common/SCALE13/...		

TurkOpticon plugin

Workers rate Requesters based on a Requester's:

- **communicativity**: How responsive has this requester been to communications or concerns you have raised?
- **generosity**: How well has this requester paid for the amount of time their HITs take?
- **fairness**: How fair has this requester been in approving or rejecting your work?
- **promptness**: How promptly has this requester approved your work and paid?

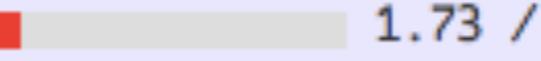
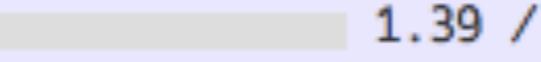
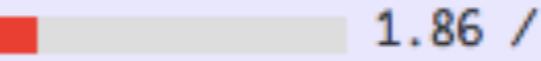
Choose the best category for this government project (good english important)

Requester:

 [The Public Group](#)

HIT Expiration Date:

Sep 10, 2013 (6 days)

communicativity:			1.17 / 5
generosity :			1.73 / 5
fairness :			1.39 / 5
promptness :			1.86 / 5

Time Allotted:

60 minutes

[What do these scores mean?](#)

Scores based on [81 reviews](#)

Terms of Service violation flags: 0

[Report your experience with this requester »](#)

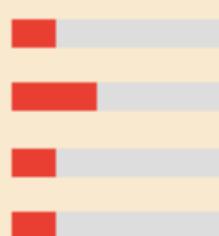
[Contact Us](#) | [Careers at Amazon](#) | [Developers](#) |

©2005-2013 Amazon.com, Inc. or its Affiliates



Took a leap of faith on this requester and was rewarded with a %50 reject rate and a broken search feature and no feedback. Would not recommend, even if you have thousands of HITs under your belt to cushion the inevitable rejections.

Aug 29 2013 | [KBH19](#) | flag | comment



Arbitrarily rejected over half of the hits I submitted, and then banned me from submitting any more hits for them. I suppose that's a blessing in disguise though, as I had no intention of doing any for them again after the first batch of rejections.

Aug 21 2013 | [bour...@g...](#) | flag | comment



Their HIT is very unclear. There is an option to browse for the result, but it does not work.

Aug 20 2013 | [jeff...@g...](#) | flag | comment

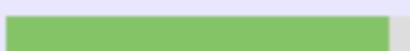
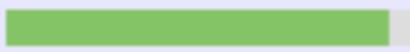
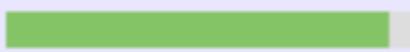
Word Alignment

Requester:

 [Chris Callison-Burch](#)

HIT Expiration Date:

Nov 12, 2013 (9 weeks)

communicativity:	 4.05 / 5
generosity :	 4.01 / 5
fairness :	 4.25 / 5
promptness :	 4.01 / 5

[What do these scores mean?](#)

Scores based on [95 reviews](#)

Terms of Service violation flags: 1

[Report your experience with this requester »](#)

Time Allotted:

60 minutes

[Contact Us](#) | [Careers at Amazon](#) | [Developers](#) |

©2005-2013 Amazon.com, Inc. or its Affiliates

For those of you who know Arabic, this is a very solid requester with a very fair pay. Highly recommended for those who want to make some real money. Payment usually take one week depending on the HITs you are doing

don't waste your time trying to submit machine translated crap, or random answers to multiple choice questions, you will get blocked instantly

Jul 27 2013 | [hala...@h...](#) | [flag](#) | [comment](#)

 Chris is one of the better requesters on MTurk, if you meet his qualifications and actually do the work as he requires. Glad to see that someone out there can finally work on those Arabic translation HITs that we've all seen for months now.

Jul 27 2013 [baudelai...@m...](#)

Good requester. Everything approved in a couple of days. I had no problems. This is a safe requester to work for.

Turker Nation discussion boards

- A watering hole for Turkers to discuss MTurk and Requesters
- Has a Requester Hall of Fame / Shame
- Lots of engaging conversation
- Sometime people vent their frustration

Here's an excerpt from an IRB application Chris Callison-Burch posted - "We will pay participants small sums of money to complete our tasks, ranging from \$0.01 to \$1. All participants can choose for themselves whether the compensation is fair, and opt not to do it if they deem the compensation to be too low. Amazon's Mechanical Turk has many other researchers and companies offering tasks, so we will offer compensation that is similar to what others offer."

He first refers to Amazon's Mechanical Turk as "an online labor market." And that, I agree with. It is an online labor market.

Requesters like him, and Dolores Labs, **collude**, explicitly or implicitly, **to keep wages** at a sub compatible with existence standard.

Unlike Jewels, I don't blame workers for taking low paying jobs. I can't blame a person for being needy enough to take what amounts to a crust of bread. **I blame Chris Callison-Burch, and others like him, for keeping the standard wage at crust of bread level.**

I feel like Maria in "Metropolis."

I tried one of those to see, I gave it up at 4 minutes in and about 2/3 of the way through. For the whole hit, I'd have taken about 6 minutes. 10 hits an hour - **\$1.70 an hour**. Restricted to U.S. residents.

This is far too low to be considered a fair wage for a U.S. resident. My performance may be very far off from what others can do. Perhaps I took 4 times or more as long as an average worker would.

My complaint is that any U.S. requester knows what wage rate is required for a U.S. resident to survive. We may not agree on an exact number. But as they say, I know a fair wage when I see it, and this is not it.

Mturk is actually much smaller than what it can appear to be. Something close to requester monopoly has the power to keep wages low. Requester co-operation, explicit or implicit, reinforces this.

Chris Callison-Burch is not unaware, I think, of the mechanics of the wage structure of Mturk.

**WORKERS
OF THE WORLD
UNITE!**



crowd-workers.com

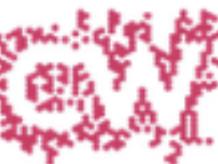
- I am developing a browser plug-in that will improve the MTurk UX for Workers
- The idea is to track and aggregate statistics across many workers, so that they have better information
- Academically, I am interested in these questions
 - a) How much time does the average Turker spend working?
 - b) What is their hourly rate?
 - c) How reputable are requesters (what fraction of the HITs do they approve v. reject)?
 - d) How much time to workers spend searching v working?

qualitative v quantitative

TurkOpticon's qualitative attributes	CrowdWorker's quantitative equivalents
promptness: How promptly has this requester approved your work and paid?	Expected time to payment: On average, how much time elapses between submitting work to this Requester and receiving payment?
generosity: How well has this requester paid for the amount of time their HITs take?	Average hourly rate: What is the average hourly rate that other Turker make when they do this requester's HITs?
fairness: How fair has this requester been in approving or rejecting your work?	Approval/rejection rates: What percent of assignments does this Requester approve? What percent of first-time Workers get any work rejected?
communicativity: How responsive has this requester been to communications or concerns you have raised?	Reasons for rejection: Archive of all of the reasons for Workers being rejected or blocked by this Requester.

CrowdWorkers Amazon Mechanical Turk

crowd-workers.com/discover?size=36

 Crowd Workers

Discover Track Your Stats

Chris Callison
ID: A23KO27

Search for HIT Minimum hourly rate Rate Search

Discover

Here are HITs that other Crowd Workers completed. You can sort by the hourly rate averaged across all workers, and other features.

Hitgroup name, description and keywords	Requester	HITs available	Hourly rate	Reward\$
Parents of 11- to 18-month-olds: Complete an online study with your child about intuitive probability! go to MTurk page For parents/guardians to complete with child aged 11 to 18 months ONLY: participate in a web-based experiment about how your child makes predictions. baby,babies,infant,children,child,parenting,kids,parents,causal,reasoning,probability,cognitive,development,learning,psychology,experiment	Lookit	1	\$590.55 \$2.50	
Website Content Review 6 go to MTurk page Search "warburg pincus healthcare". List the 1st option in the top navigation bar, of the 1st search result for payment. warburg,pincus,healthcare	Cam Elizabeth Harvey	1	\$29.03 \$0.03	
App Usage Study go to MTurk page Testing new apps for mobile devices over a 12-day period survey,apps,mobile	Gracy P.	1	\$25.18 \$8.00	
Website Content Review 43 go to MTurk page Search 'gramercy emerging markets high yield fund'. List the title of the 2nd option in the website navigation menu of the 1st search result for	Cam Elizabeth	1	\$23.79 \$0.03	

Demographics of MTurk

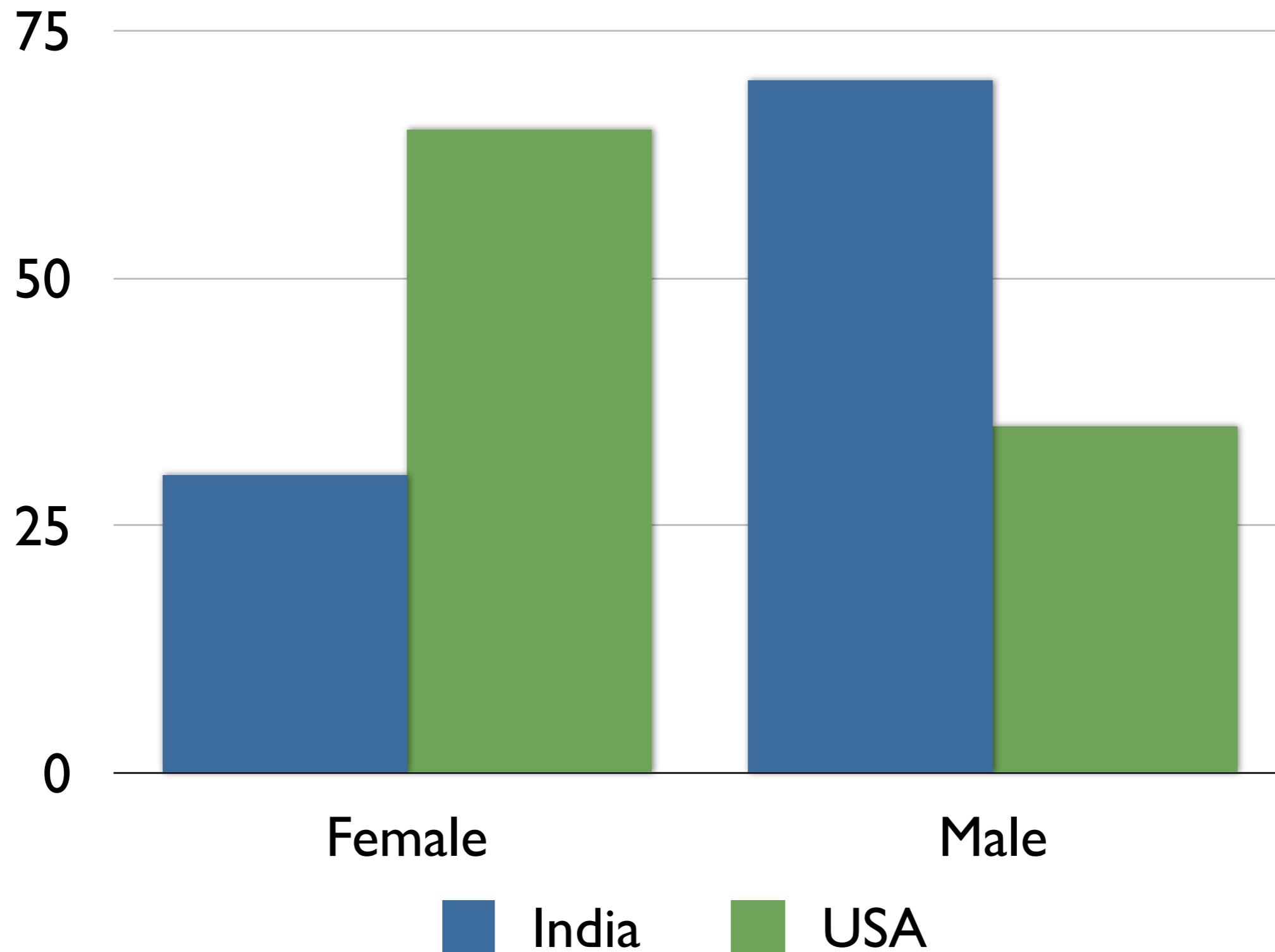
Who are the Turkers?

- Requesters are given very little information about Turkers - basically just a serial number
- No names, no demographic information (like what languages they speak)
- Who are these people who work for us?

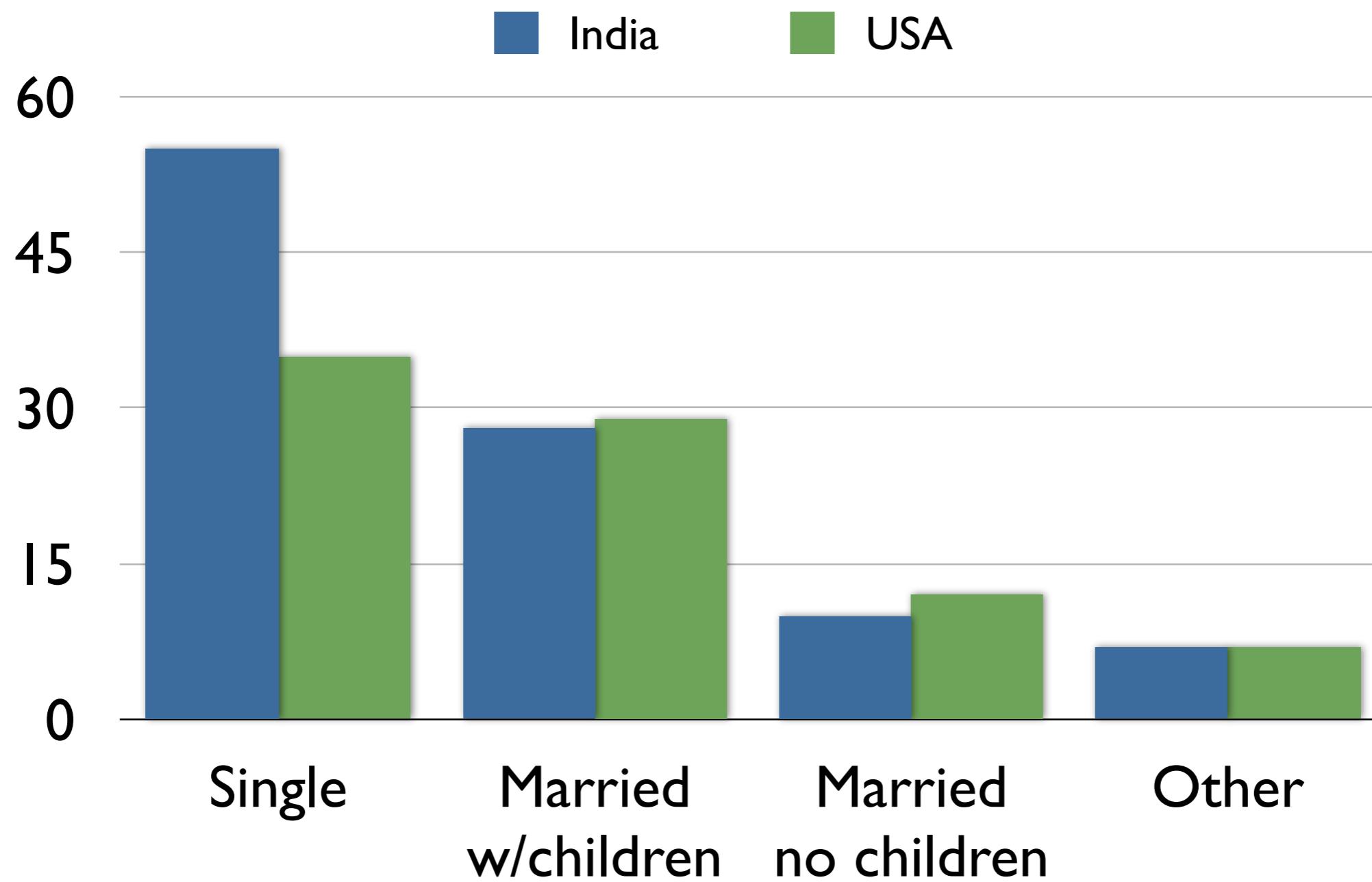
Who are the Turkers?

- Post a HIT to interview them!
- Panos Ipeirotis has a nice demographic survey on his blog “A Computer Scientist In Business School”
- I posted a 400 question political survey the week before the 2012 presidential election
- I paid 1000 Turkers \$5 releasing batches of surveys each hour at 9am in each time zone

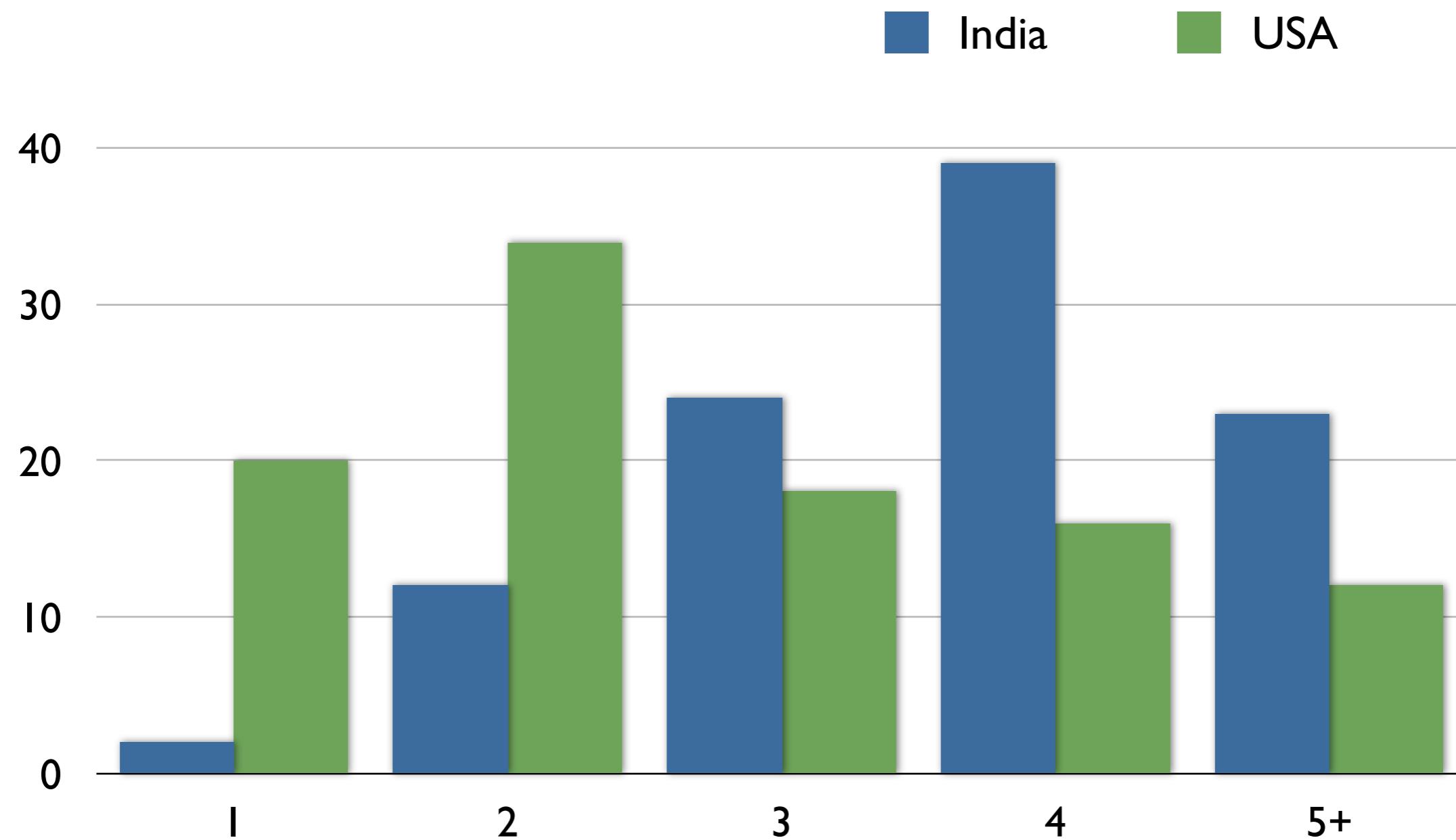
Gender breakdown



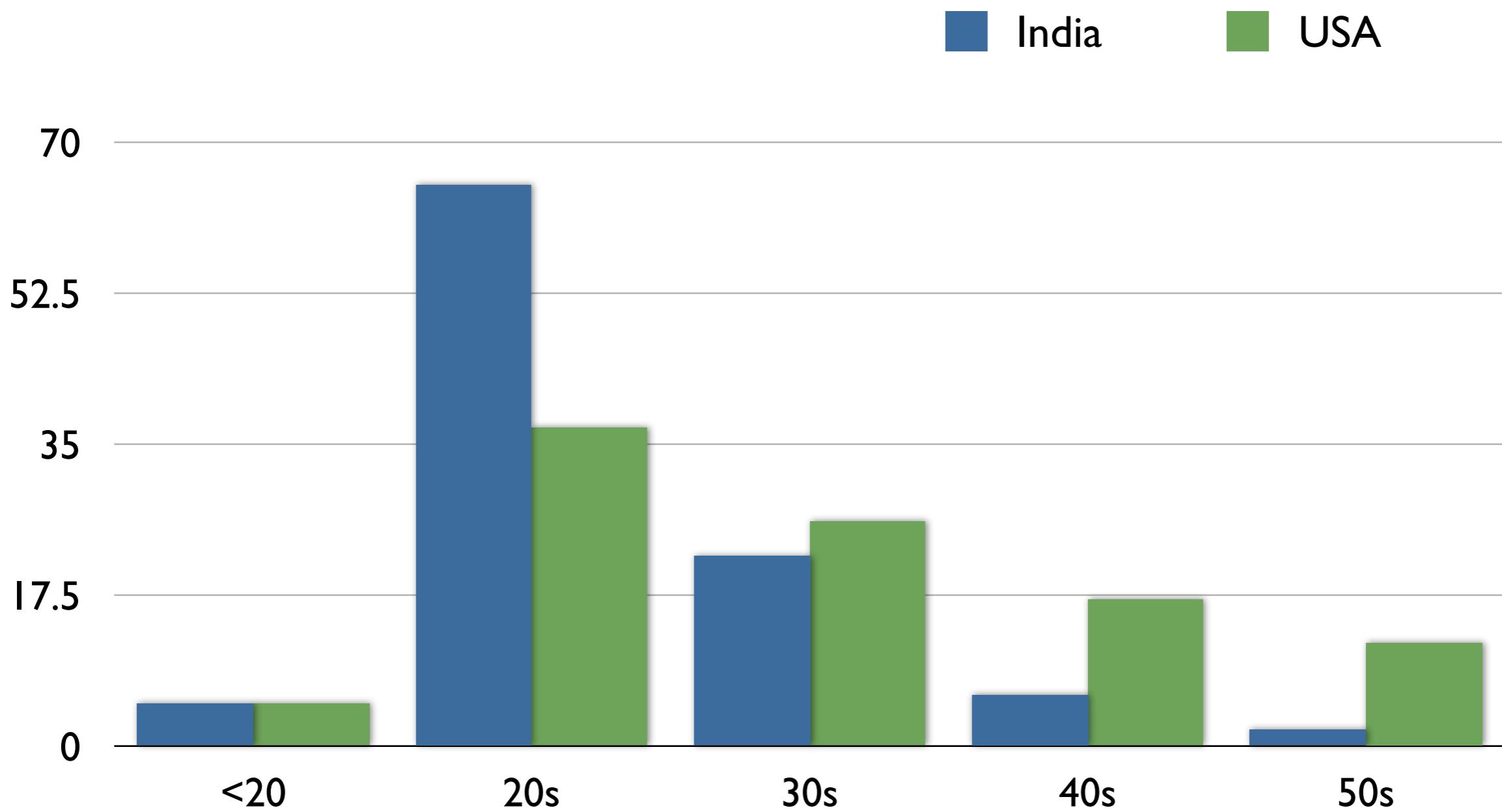
Marital status



Household sizes

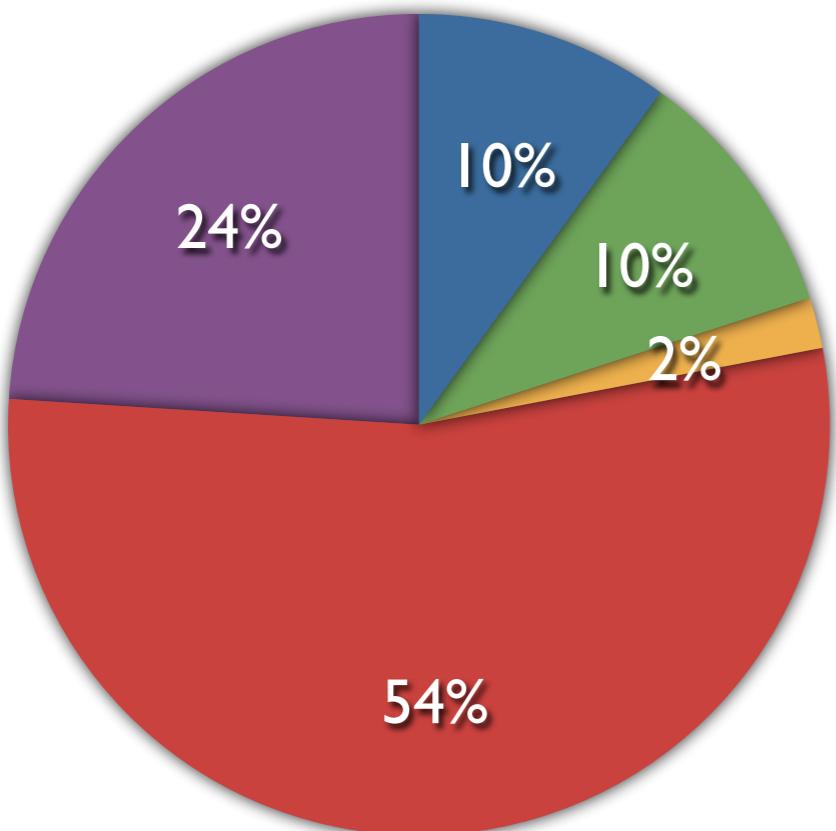


Age distribution

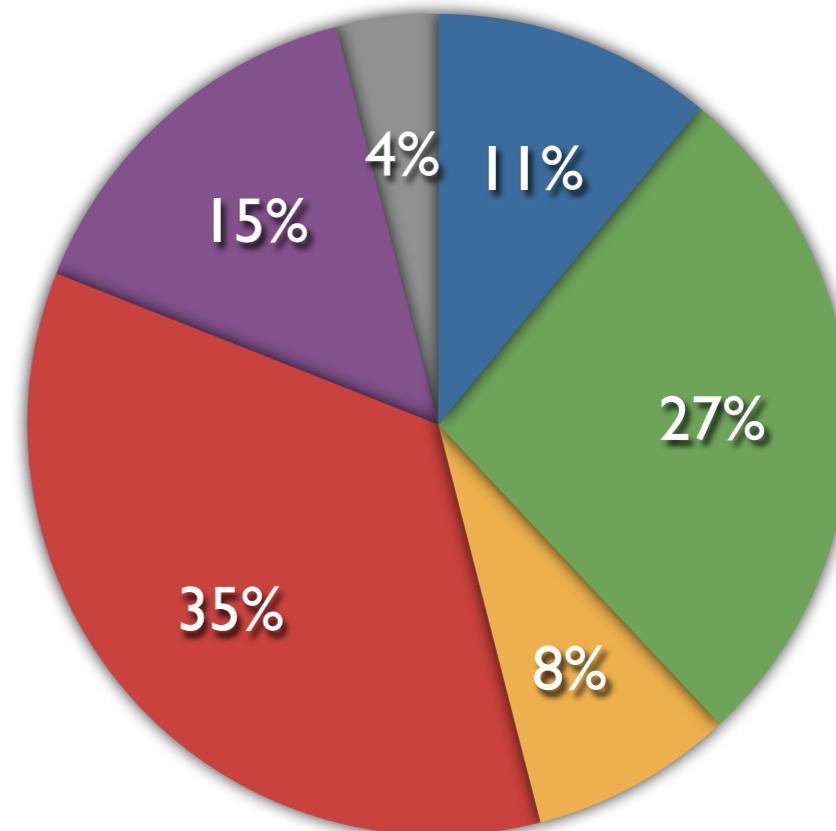


Education level

India

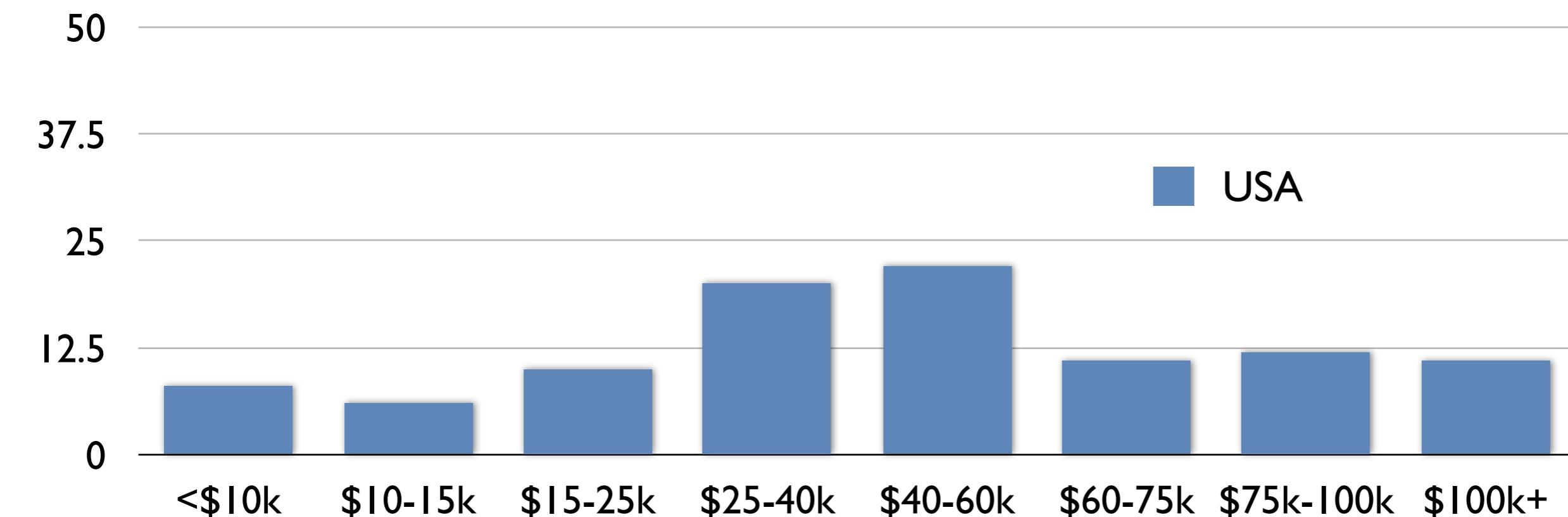
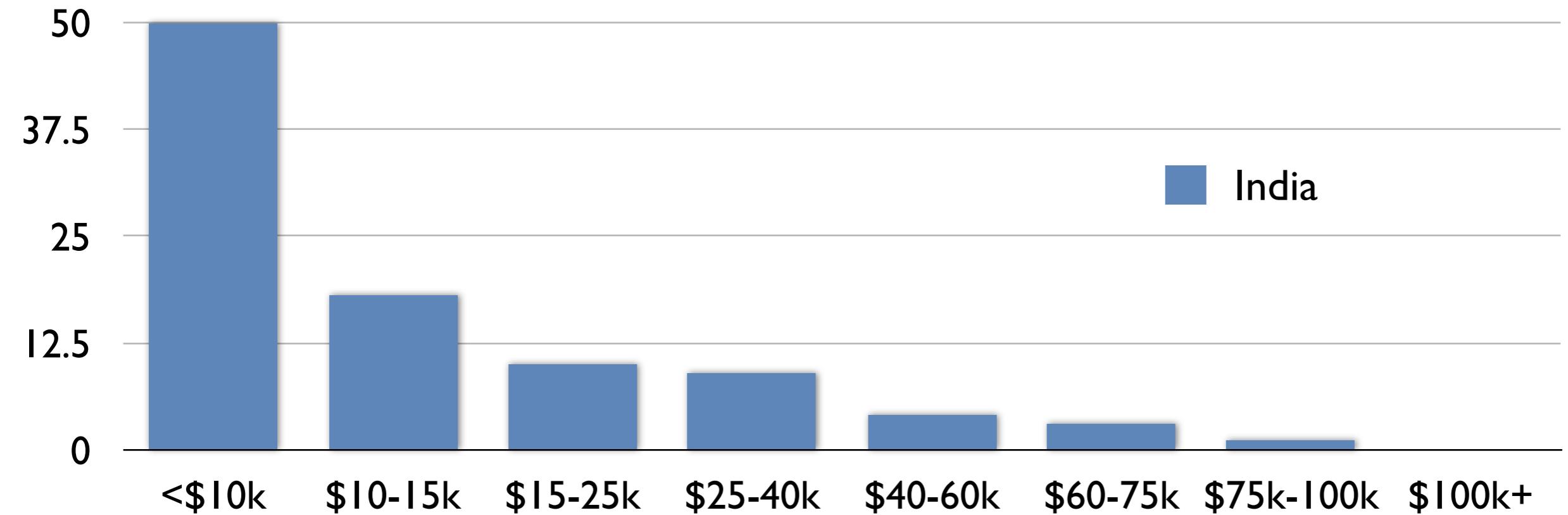


USA

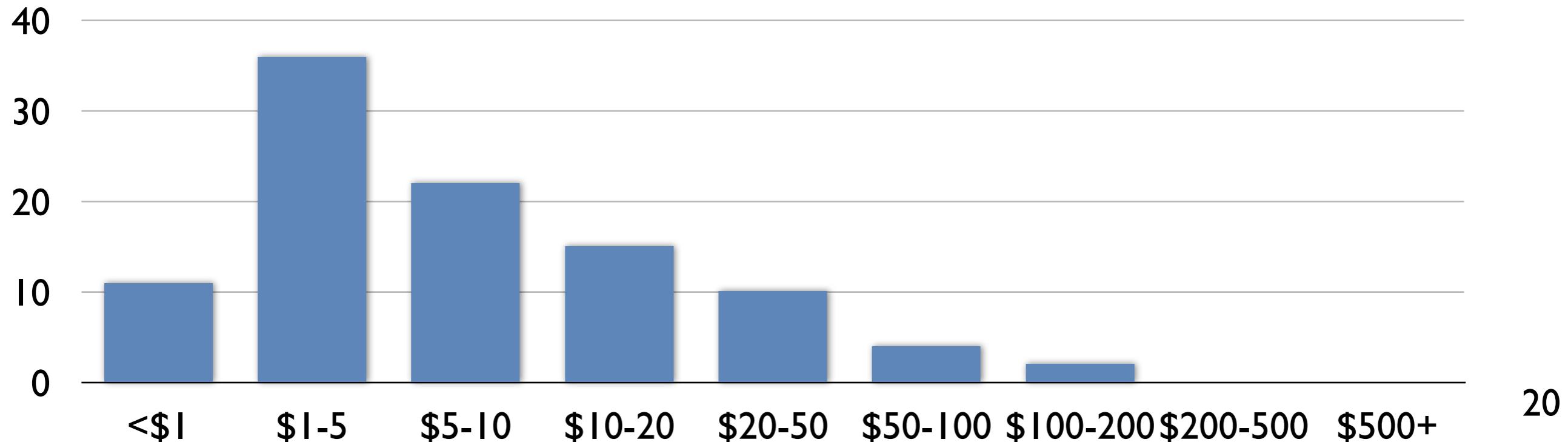
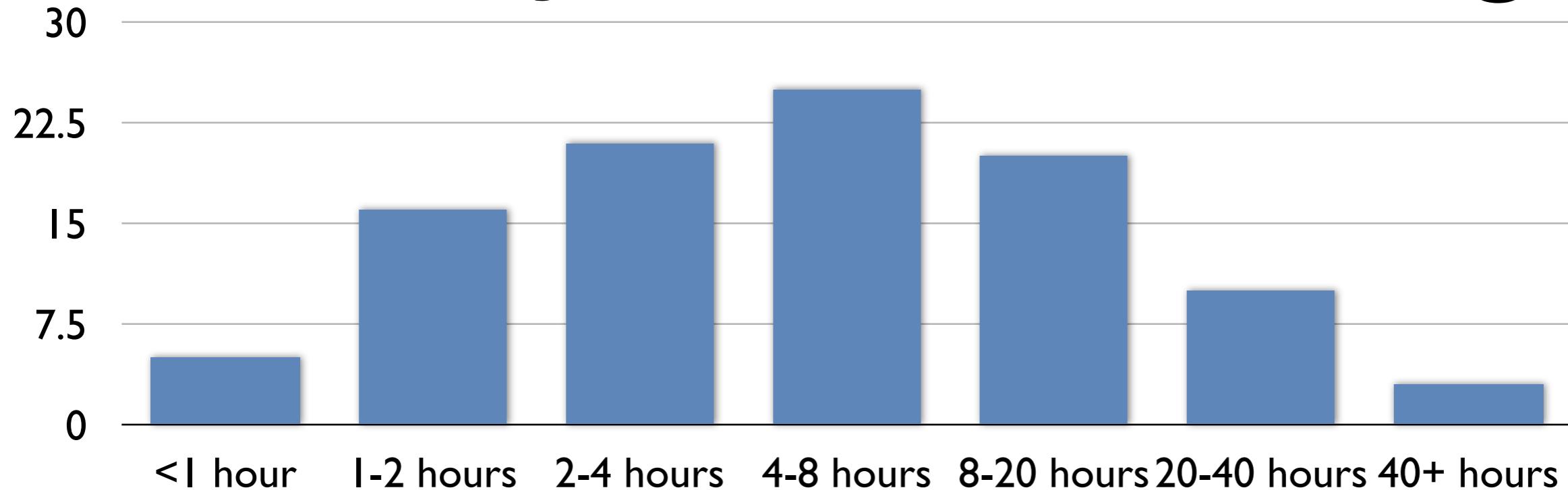


- High School
- Some college
- Associates
- Bachelors
- Masters
- PhD

Household Incomes



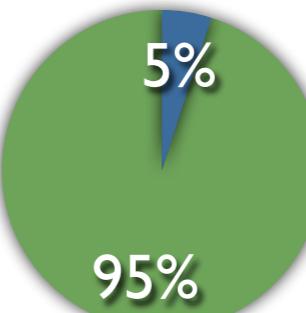
Weekly hours, earnings



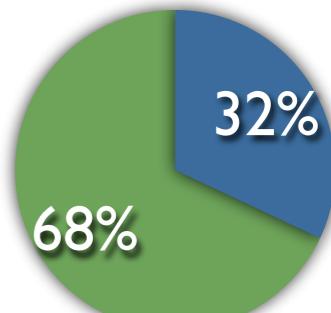
Why do you do it?

I use MTurk to kill time

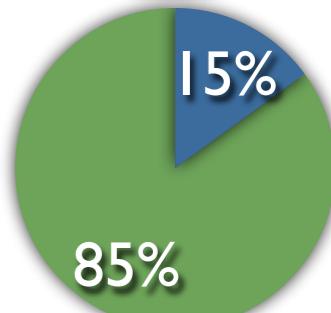
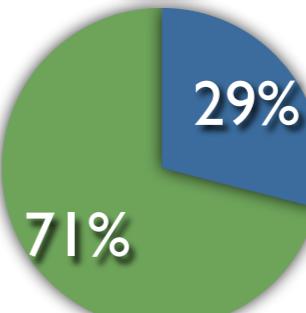
India



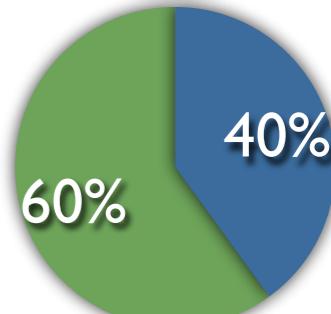
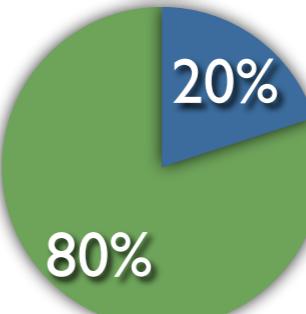
USA



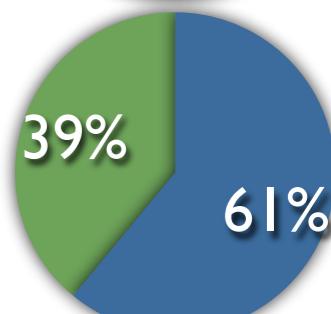
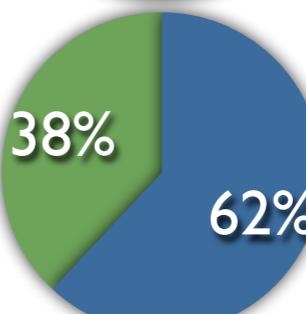
I participate on MTurk for fun



Secondary source of income
or for pocket change



Fruitful way to spend free time
and get cash (instead of TV)



● Yes

● No

● No

Survey Methodology

- Ask 1000 Turkers to fill out questionnaire
- Pay them each \$0.10 for ~3 minutes
- Do you think that the methodology was good?
- Do you think it yields an accurate sample of Turkers?
- Any bias?



American National Election Studies

<http://www.electionstudies.org>

HOME

DATA CENTER

UTILITIES

HELP CENTER

LIBRARY

What is ANES?

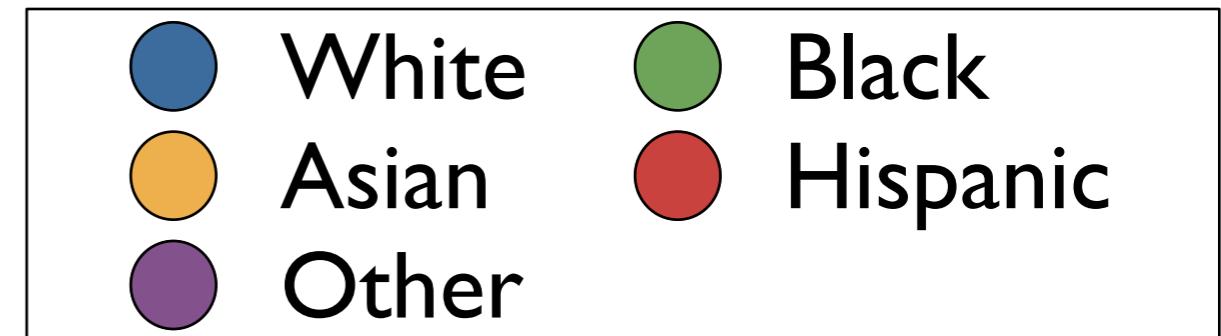
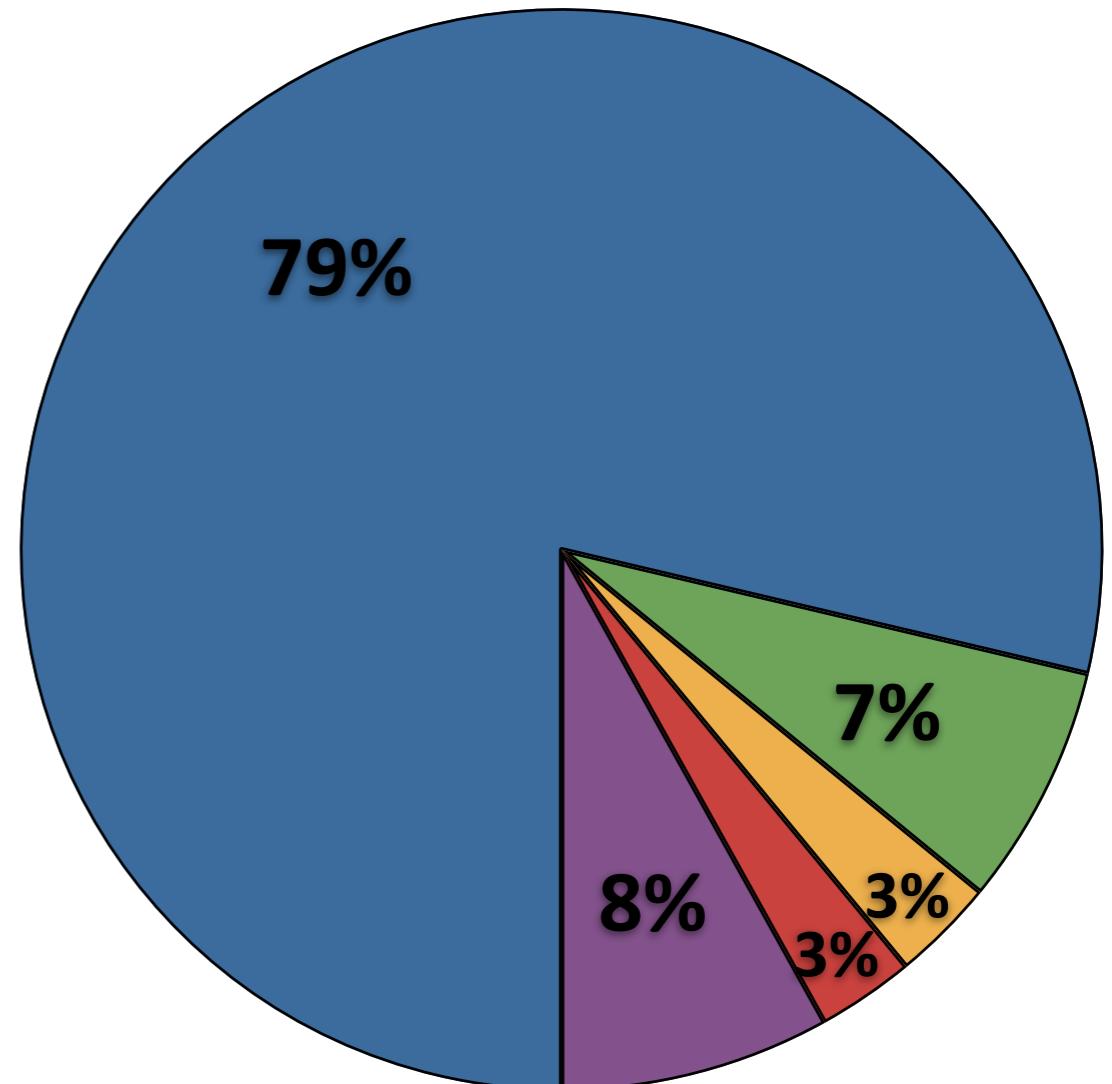
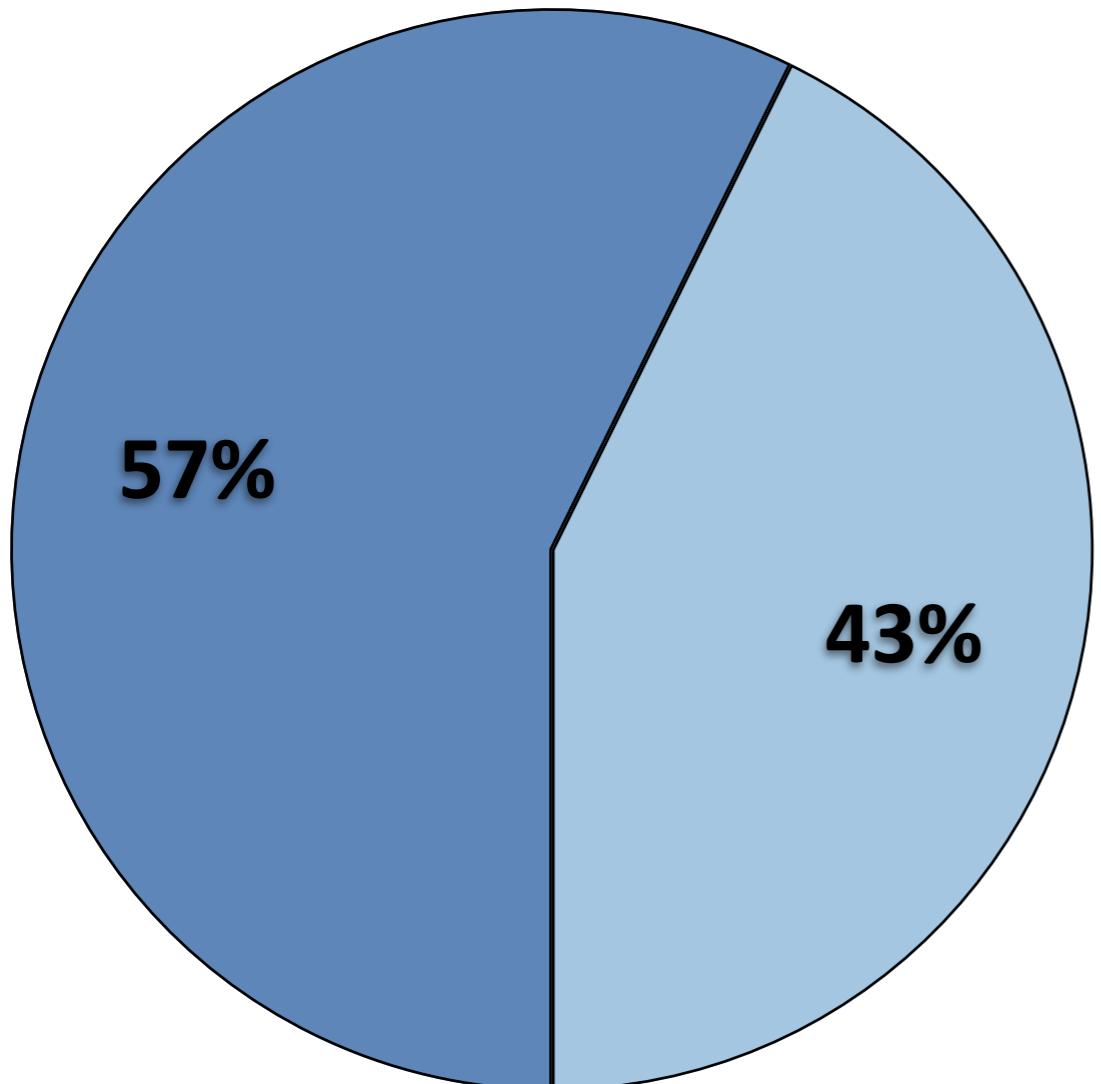
To serve the research needs of social scientists, teachers, students, policy makers and journalists, the ANES produces high quality data from its own surveys on voting, public opinion, and political participation.

Should same-sex couples be ALLOWED to marry, or do you think they should NOT BE ALLOWED to marry?

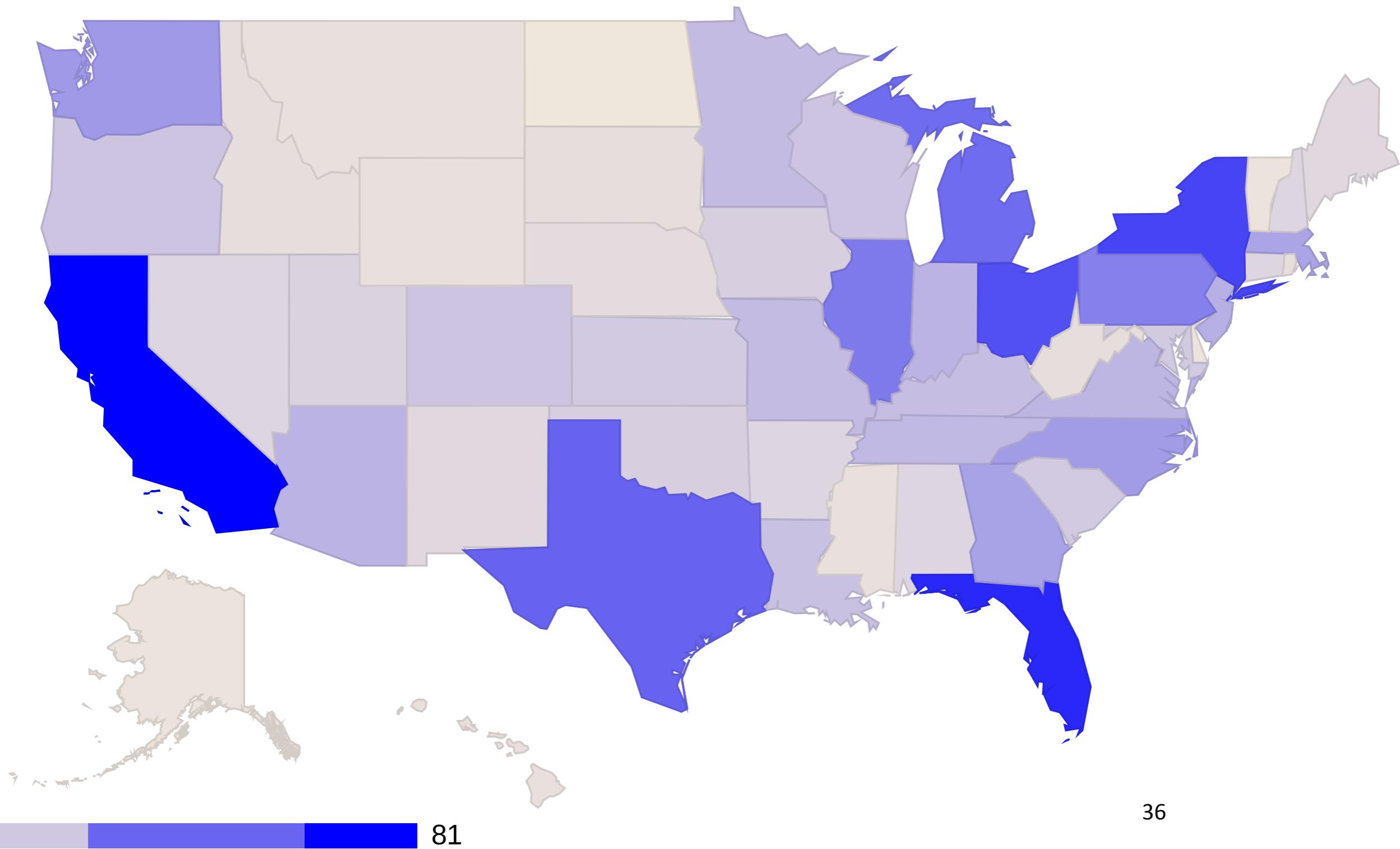
Is there something that would make you vote for a Republican presidential candidate? Please list as many reasons as you can think of.

We are interested in how people are getting along financially these days. Would you say that you / you and your family are BETTER off or WORSE off than you were a year ago?

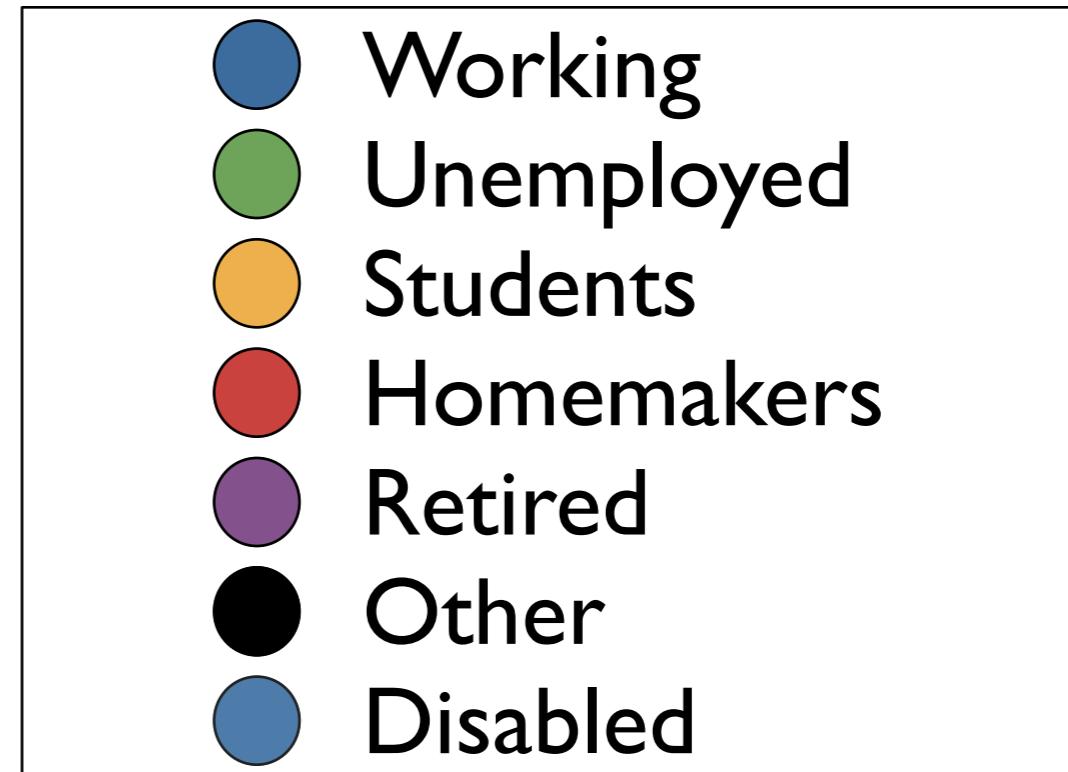
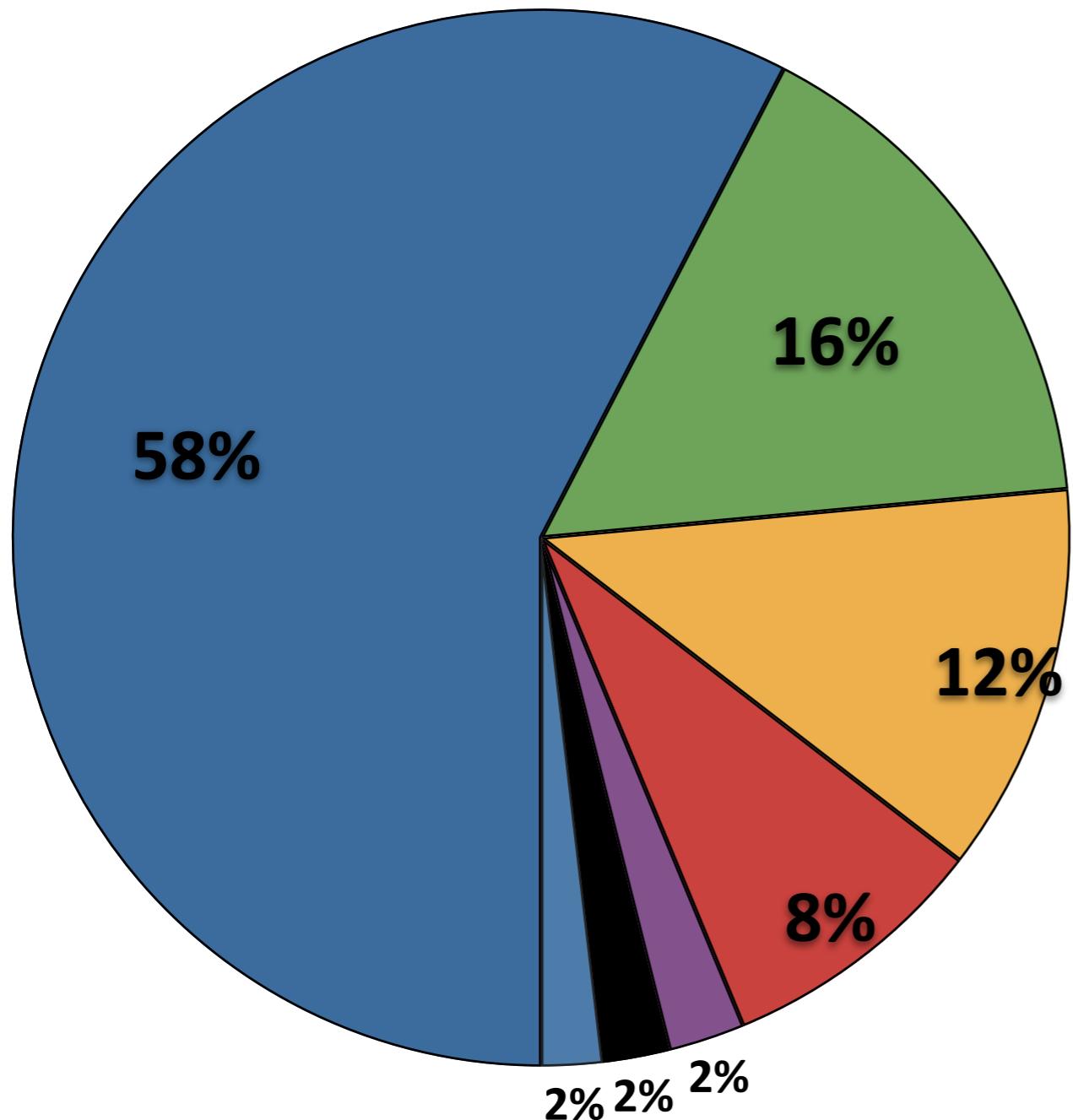
Demographics



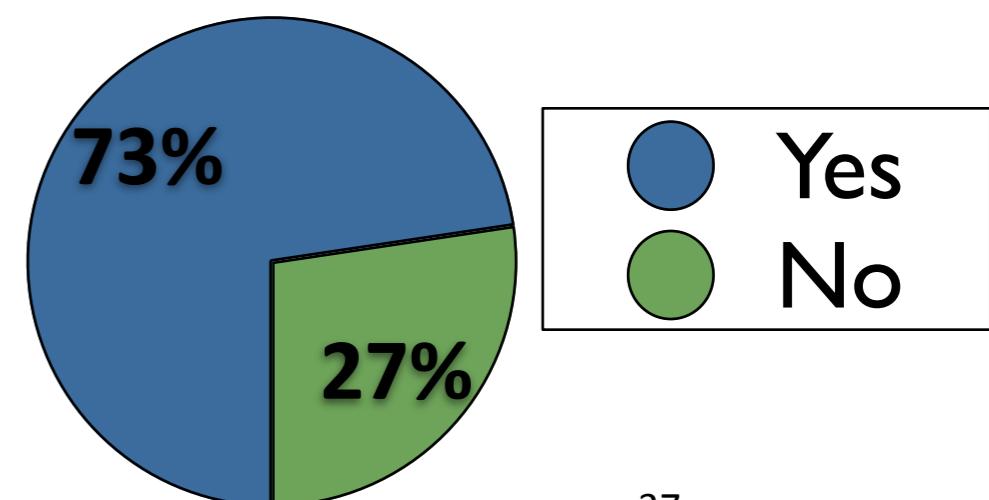
Geographic distribution



Employment status



Had a job in last 6 months?



My jobs have always consisted of customer service, cashier at a gas station, grocery store, etc. Currently, I am not employed. I stay at home and care for my disabled son.

I am an event planner for a non-profit organization. I plan events, coordinate auctions, solicit donations from corporations and individuals

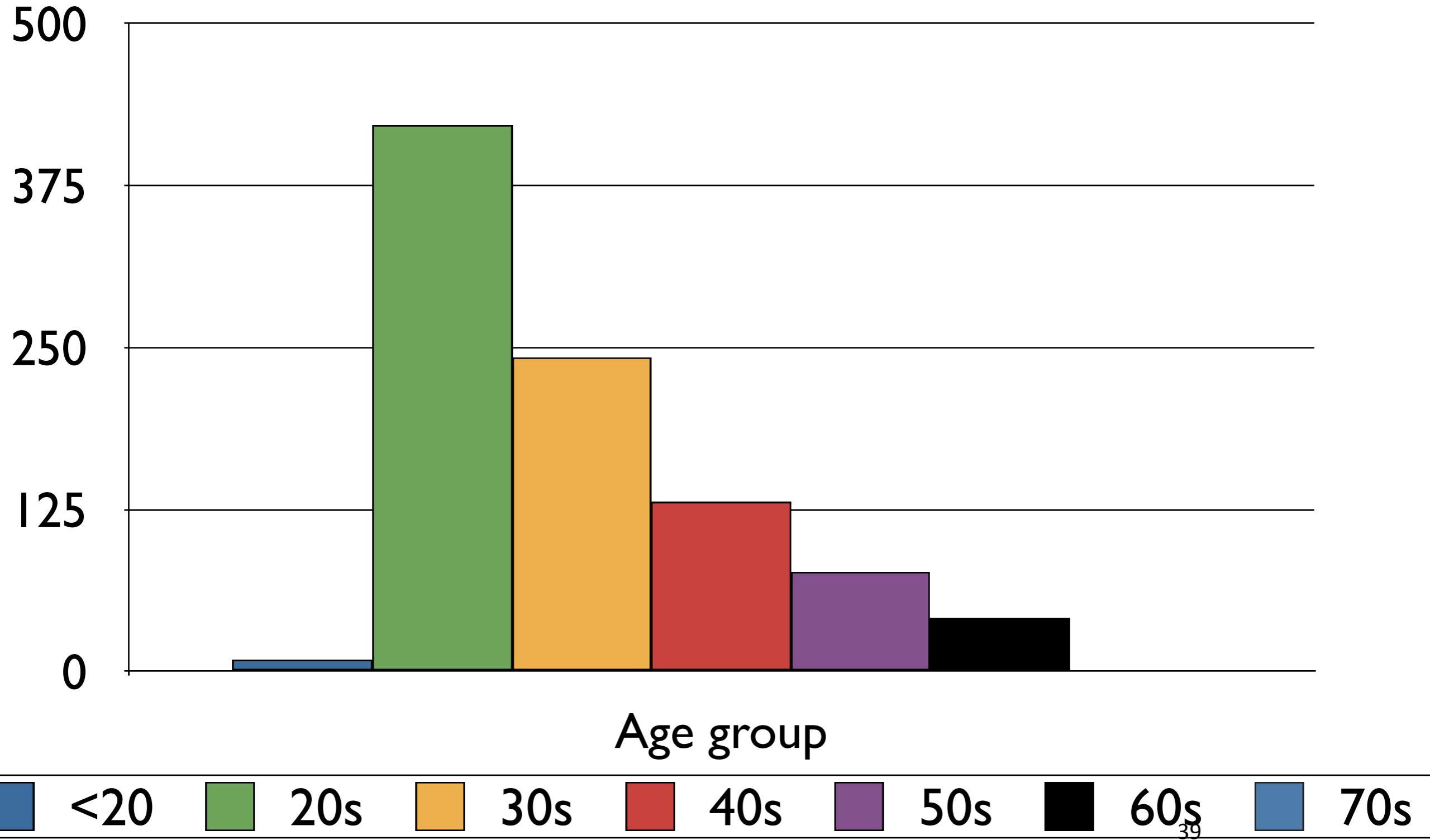
My last job was working in a subshop called Jimmy Johns. My duties were to make subs, as well as cleaning up during my shift and at close

My last regular job was for the U.S. Census Bureau. My job involved going door to door in my town and neighboring towns in my county and interviewing residents

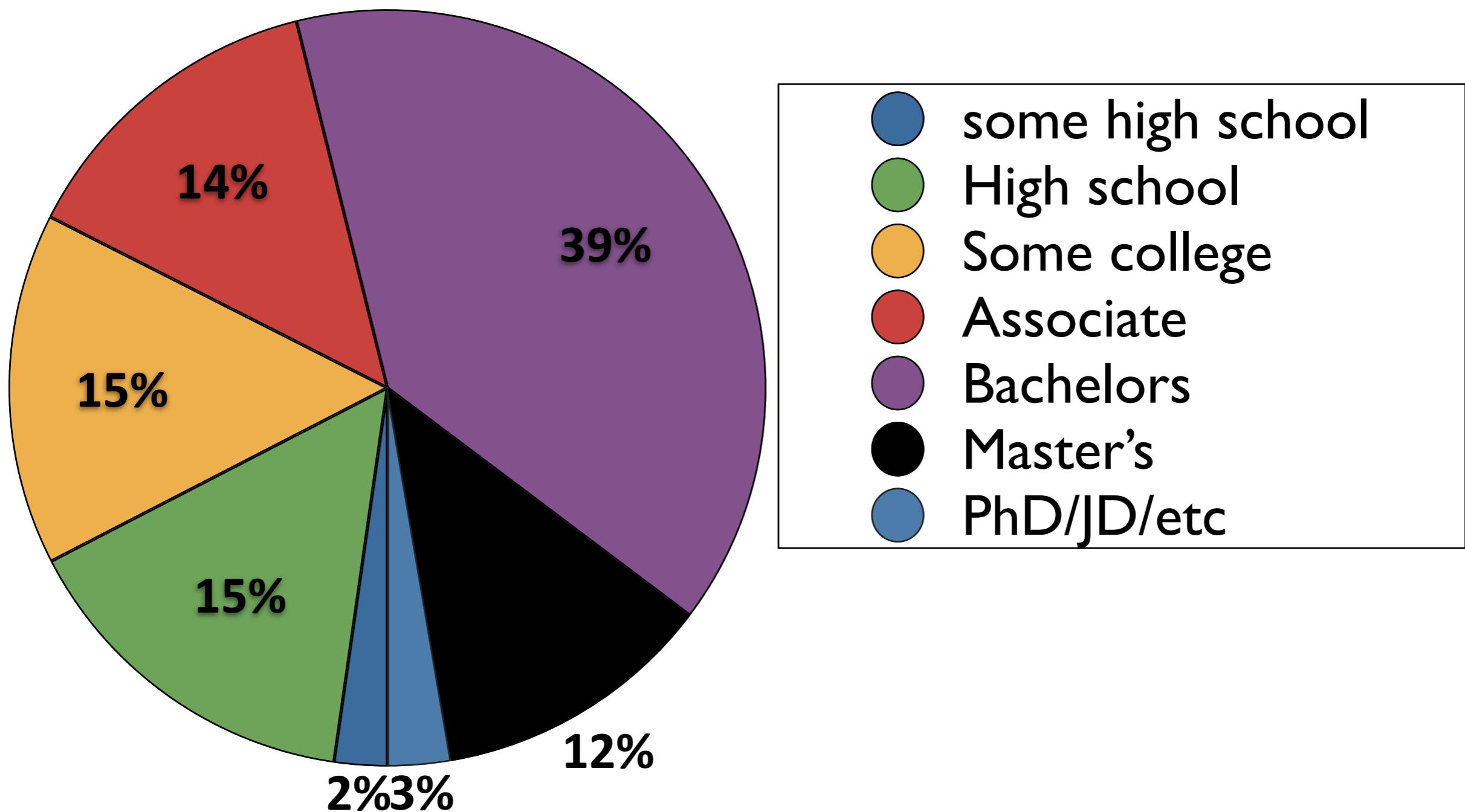
My last job was at Walgreens. I was a cashier overnight. Now, I am unemployed and I make very little money on mturk.

I am an energy efficiency analyst working at a utility. I oversee evaluation of our energy efficiency programs so that their impact can be accurately estimated

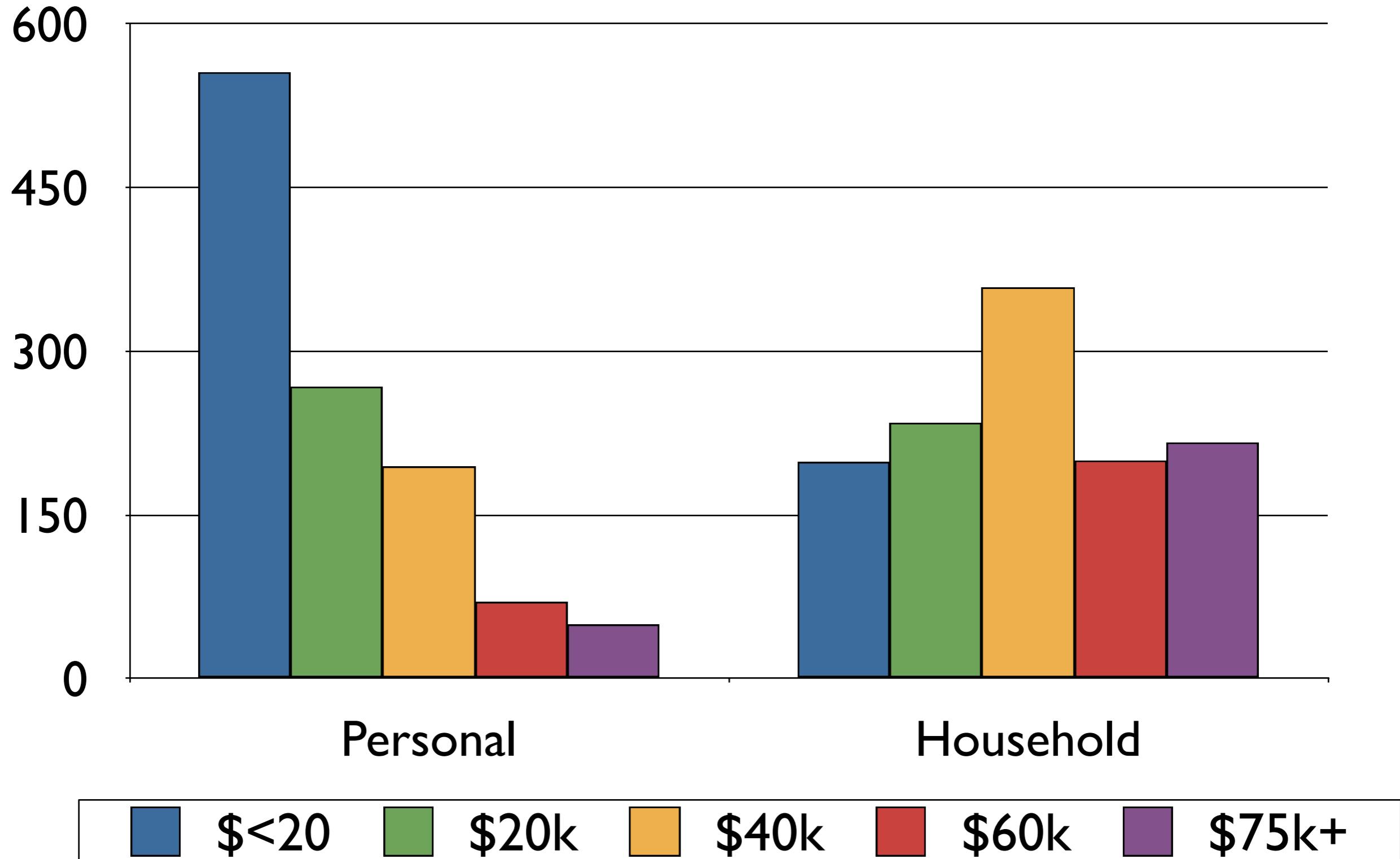
Age



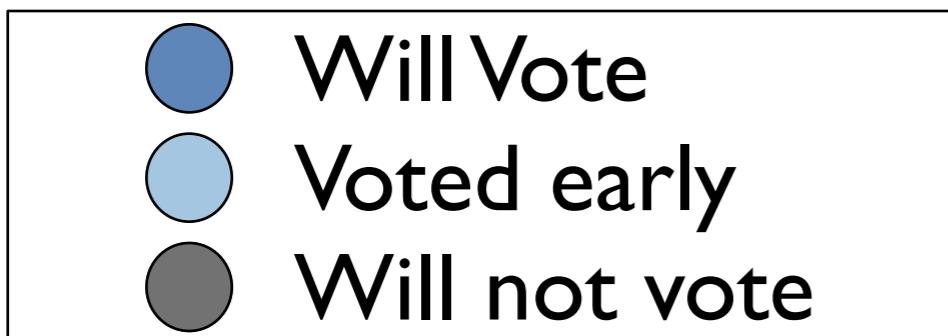
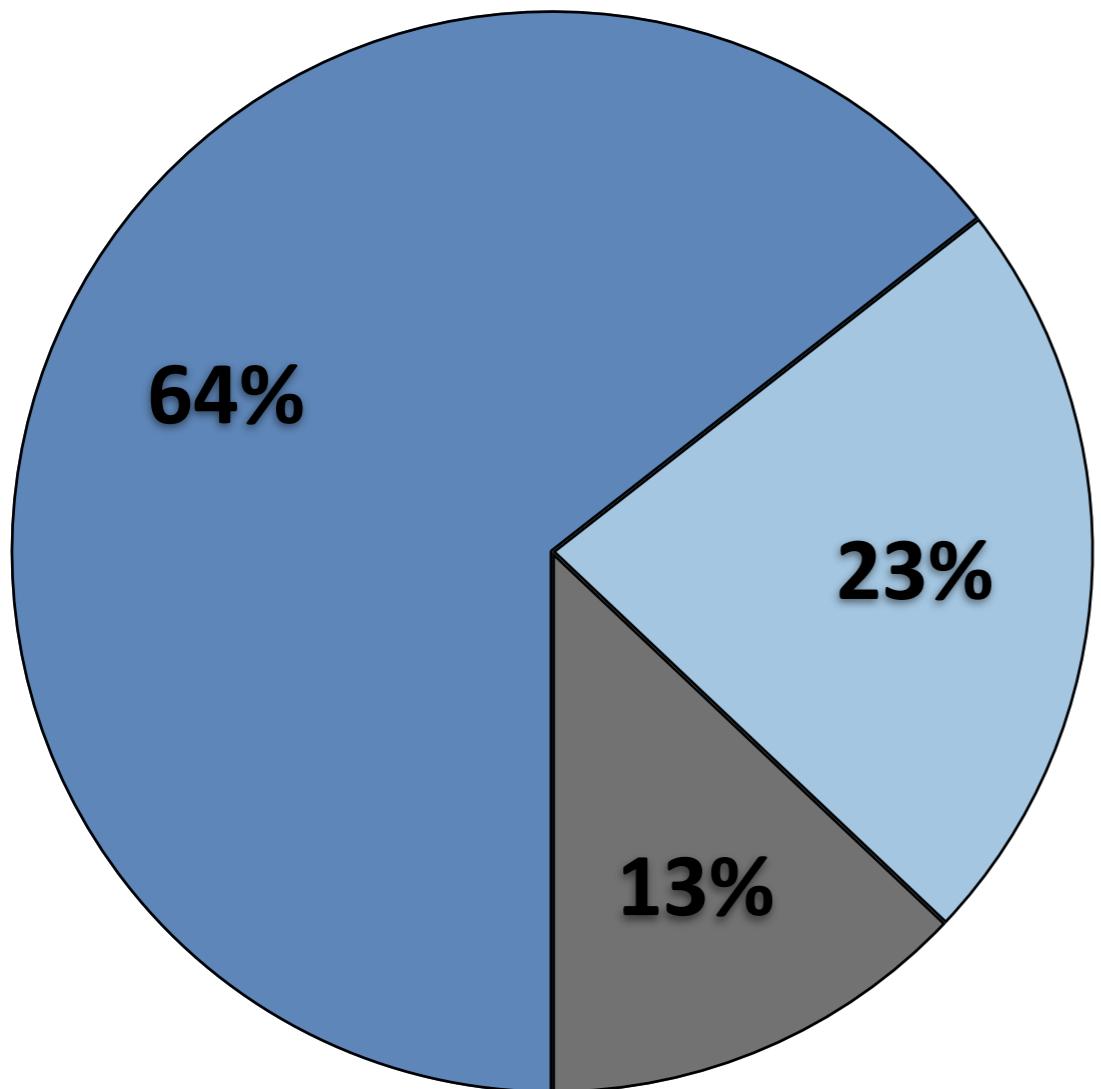
Education level



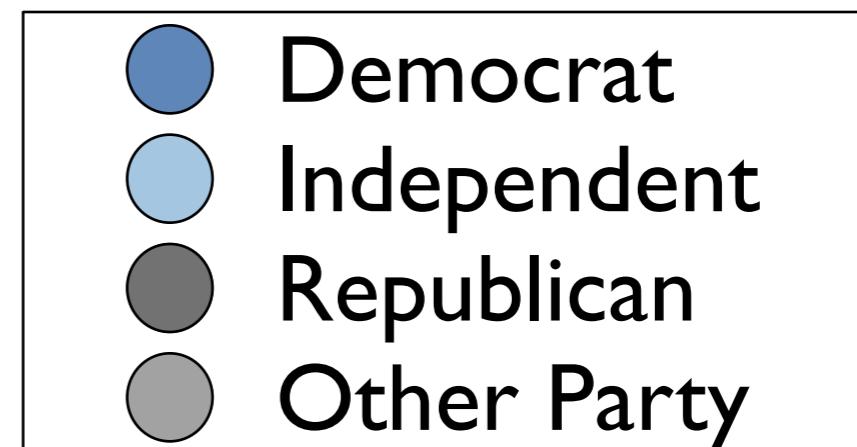
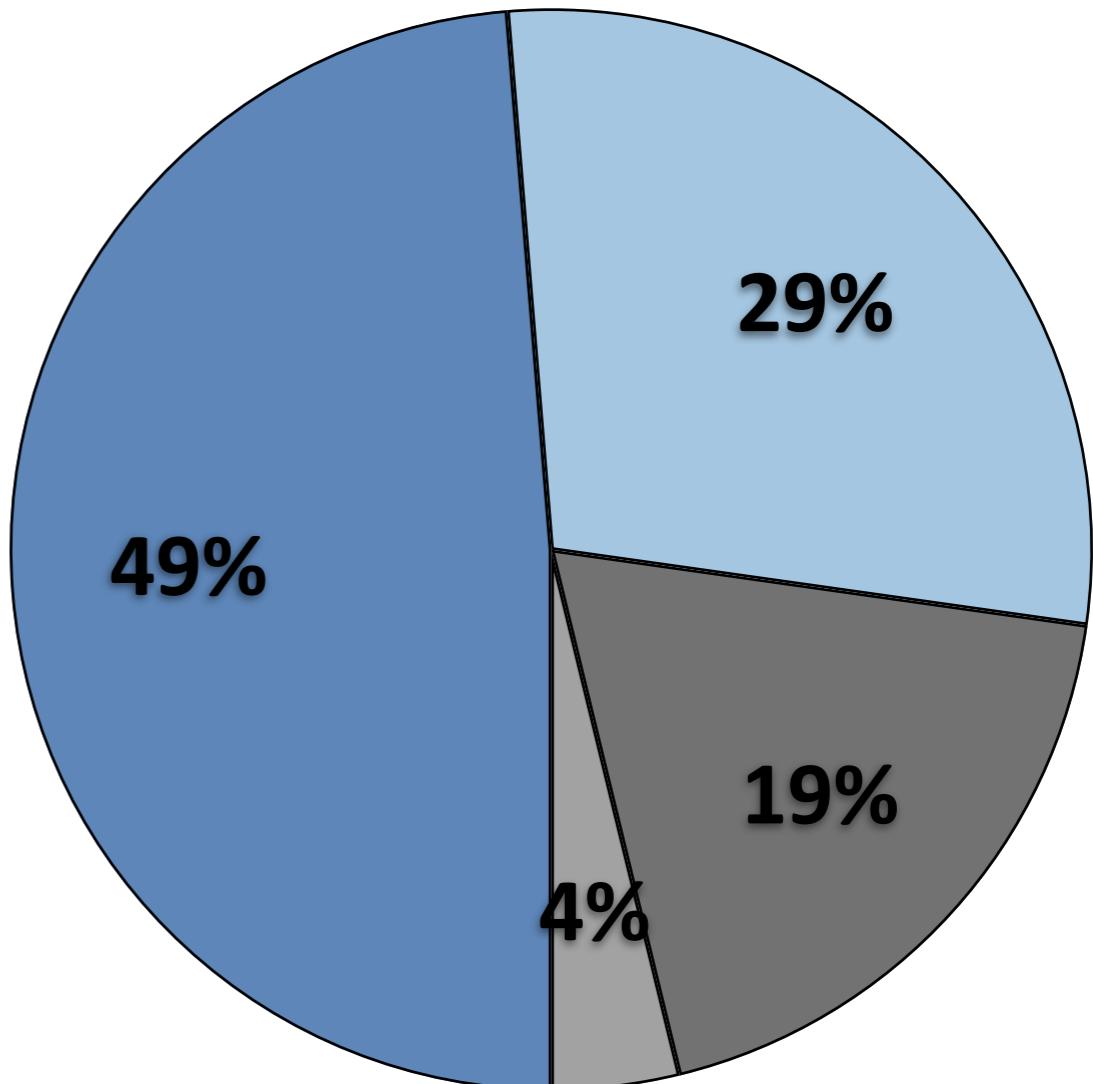
Annual Income



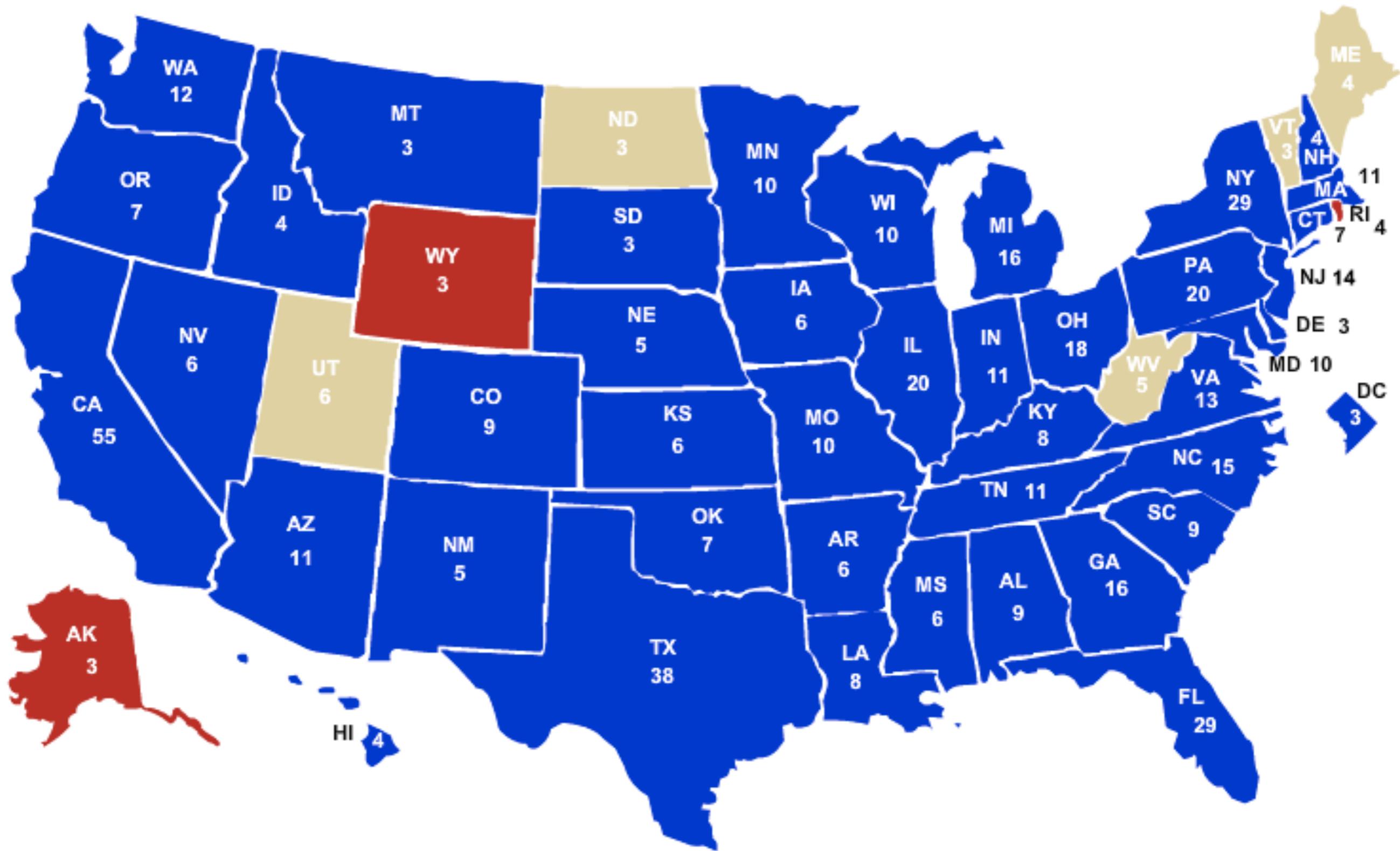
Plans to vote / Party Affiliation



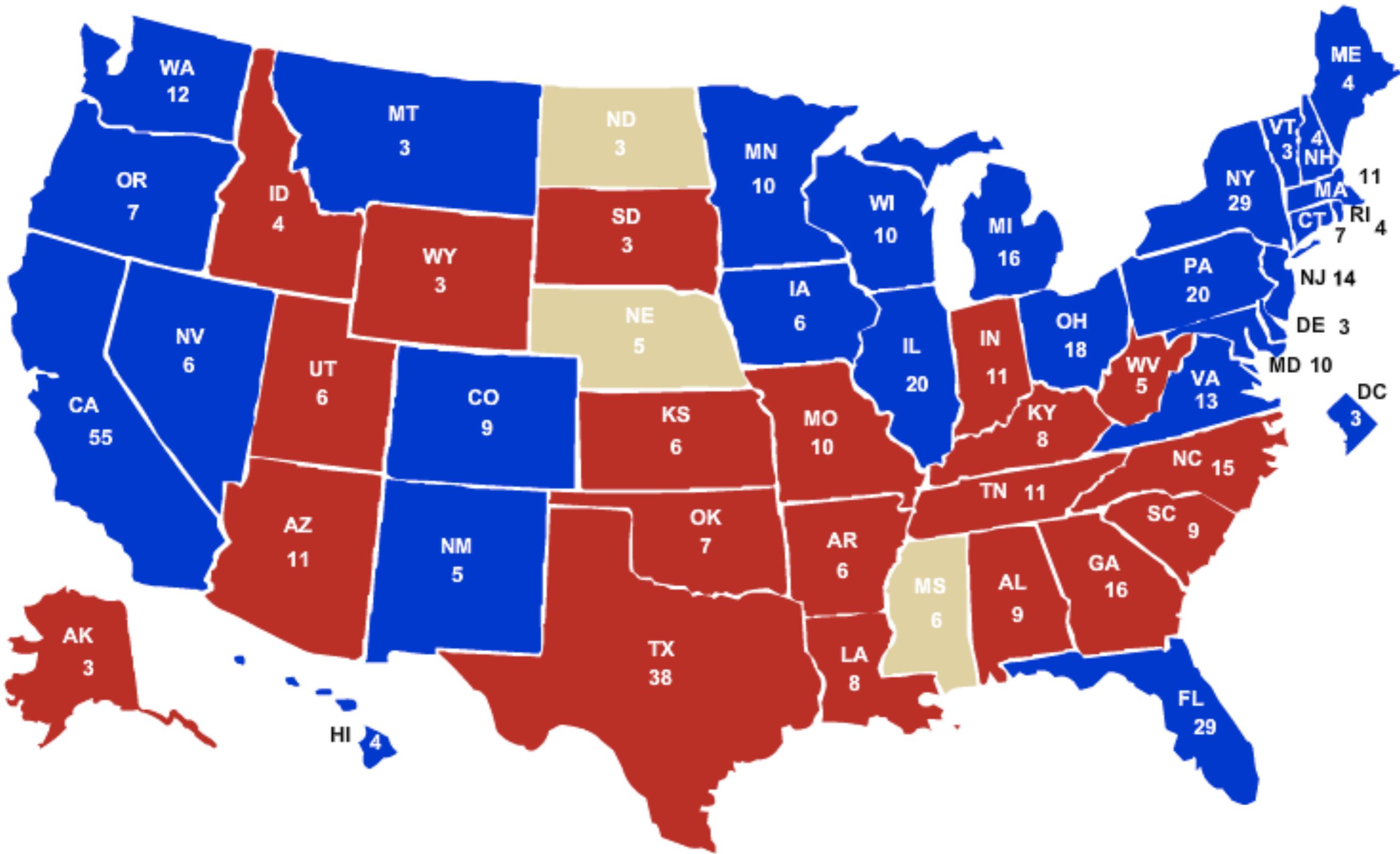
42



Obama Landslide if only Turkers voted

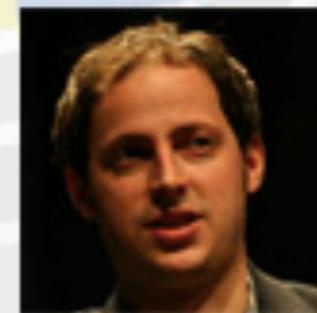
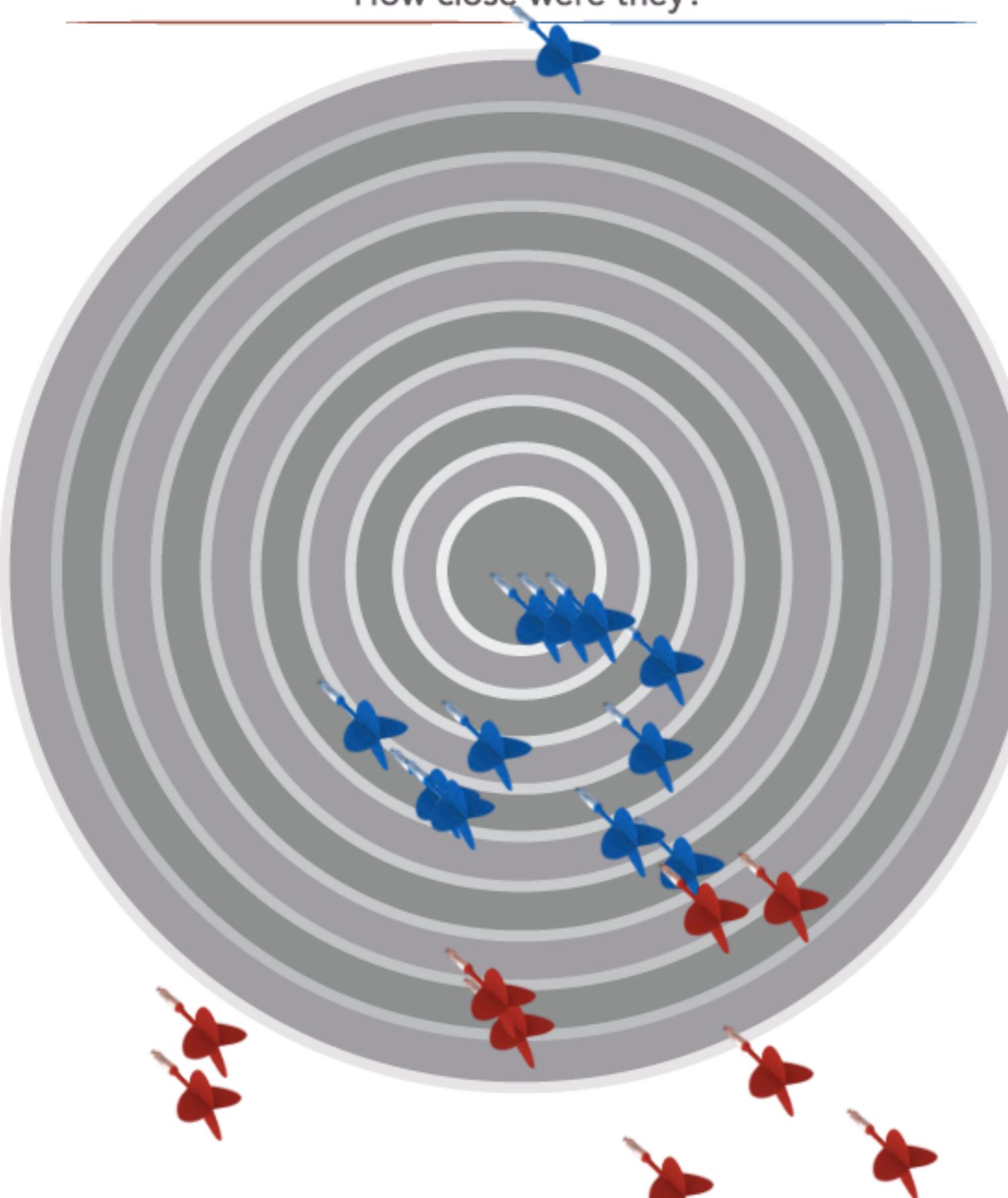


Turkers accurately predict their states



PUNDARTS

How close were they?



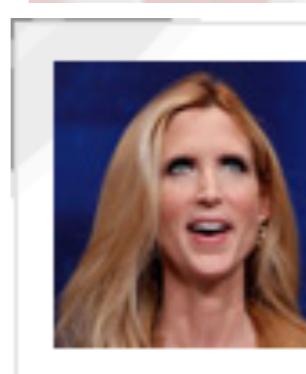
Nate Silver
The New York Times
Obama: 332
Romney: 206
Difference: 0

amazon mechanical turk
beta Artificial Artificial Intelligence

Obama: 335, Romney: 189
Error: 3, Unallocated: 14



Donna Brazile
Vice Chairwoman of the Democratic National Committee
Obama: 313
Romney: 225
Difference: -19

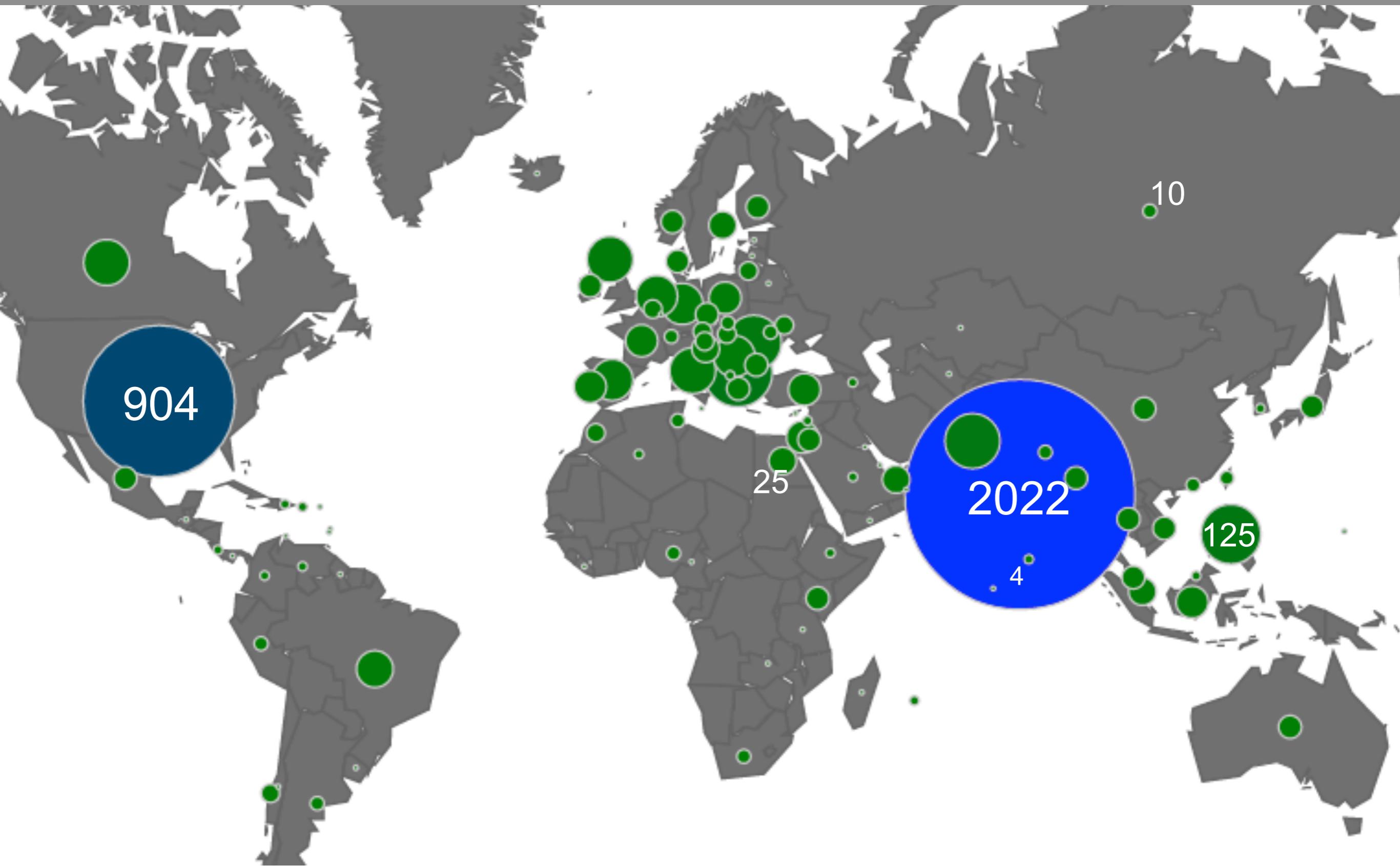


Ann Coulter
Fox News
Obama: 265
Romney: 273
Difference: -67



Karl Rove
Obama: 259
Romney: 279
Difference: -73

Worldwide Workforce



Language Demographic Study

English	591	French	63	Vietnamese	34
Tamil	253	Polish	61	Macedonian	31
Malayalam	219	Urdu	56	Cebuano	29
Hindi	149	Tagalog	54	Swedish	26
Spanish	131	Marathi	48	Bulgarian	25
Telugu	87	Russian	44	Hungarian	23
Chinese	86	Italian	43	Swahili	23
Romanian	85	Bengali	41	Thai	22
Portuguese	82	Gujarati	39	Catalan	22
Arabic	74	Hebrew	38	Lithuanian	21
Kannada	72	Dutch	37	Punjabi	21
German	66	Turkish	35	Others	≤20

How big is the MTurk
Market?

Methodology

- Crawl the list of HITs once per hour
- Record the
 - RequesterID
 - Number of HITs available
 - Reward amount
 - Title, description, keywords
 - Qualifications required

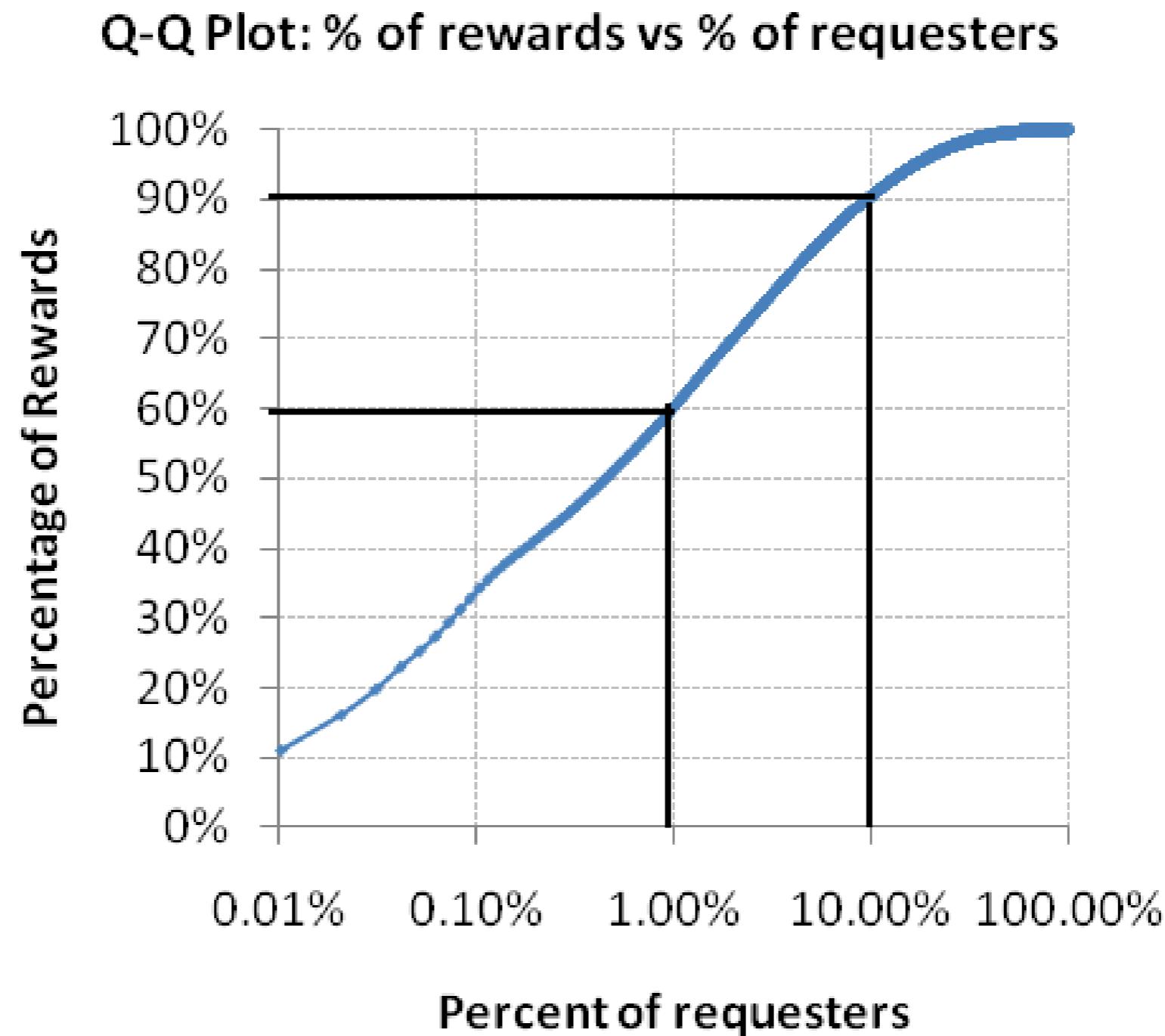
Top Requesters

Requester ID	Requester Name	#HIT groups	Total HITs	Rewards	Type of tasks
A3MI6MIUNWCR7F	CastingWords	48,934	73,621	\$59,099	Transcription
A2IR7ETVOIULZU	Dolores Labs	1,676	320,543	\$26,919	Mediator for other requesters
A2XL3J4NH6JI12	ContentGalore	1,150	23,728	\$19,375	Content generation
A11970GL0WOQ3G	Smartsheet.com Clients	1,407	181,620	\$17,086	Mediator for other requesters
AGW2H4I480ZX1	Paul Pullen	6,842	161,535	\$11,186	Content rewriting
A1CTI3ZAWTR5AZ	Classify This	228	484,369	\$9,685	Object classification
A1AQ7EJ5P7ME65	Dave	2,249	7,059	\$6,448	Transcription
AD7C0BZNKYGYV	QuestionSwami	798	10,980	\$2,867	Content generation and evaluation
AD14NALRDSN9	retaildata	113	158,206	\$2,118	Object classification
A2RFHBFTZX7UN	ContentSpooling.net	555	622	\$987	Content generation and evaluation
A1DEBE1WPE6JFO	Joel Harvey	707	707	\$899	Transcription
A29XDCTJMAE5RU	Raphael Mudge	748	2,358	\$548	Website feedback

Number of HITs and Total Value

- From January 2009 - April 2010:
 - 165,368 HIT groups
 - 6,701,406 HITs total
 - 9,436 requesters
 - The total value of the posted HITs was \$529,259

A few requesters offer most of the rewards

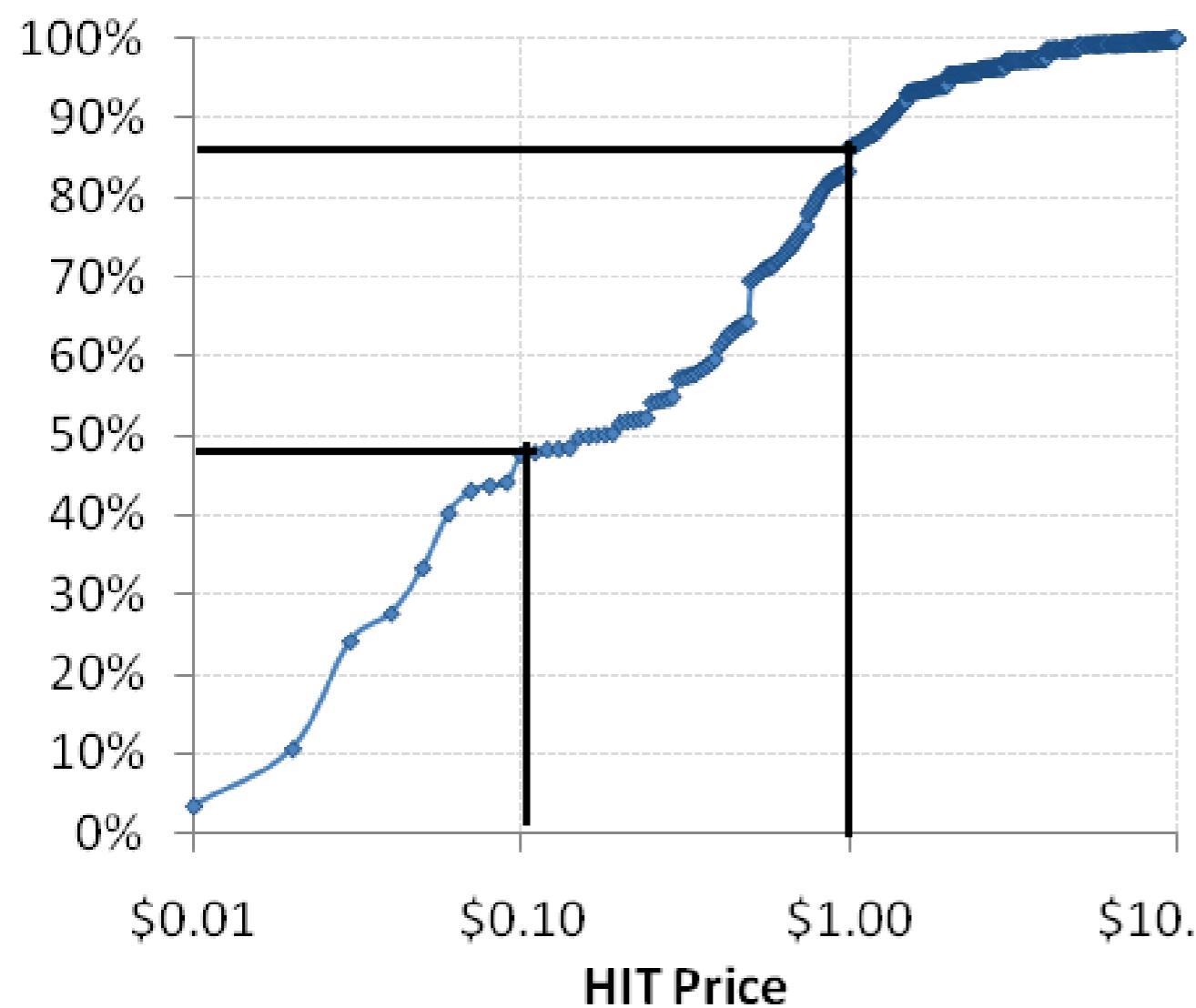


Rewards by keyword

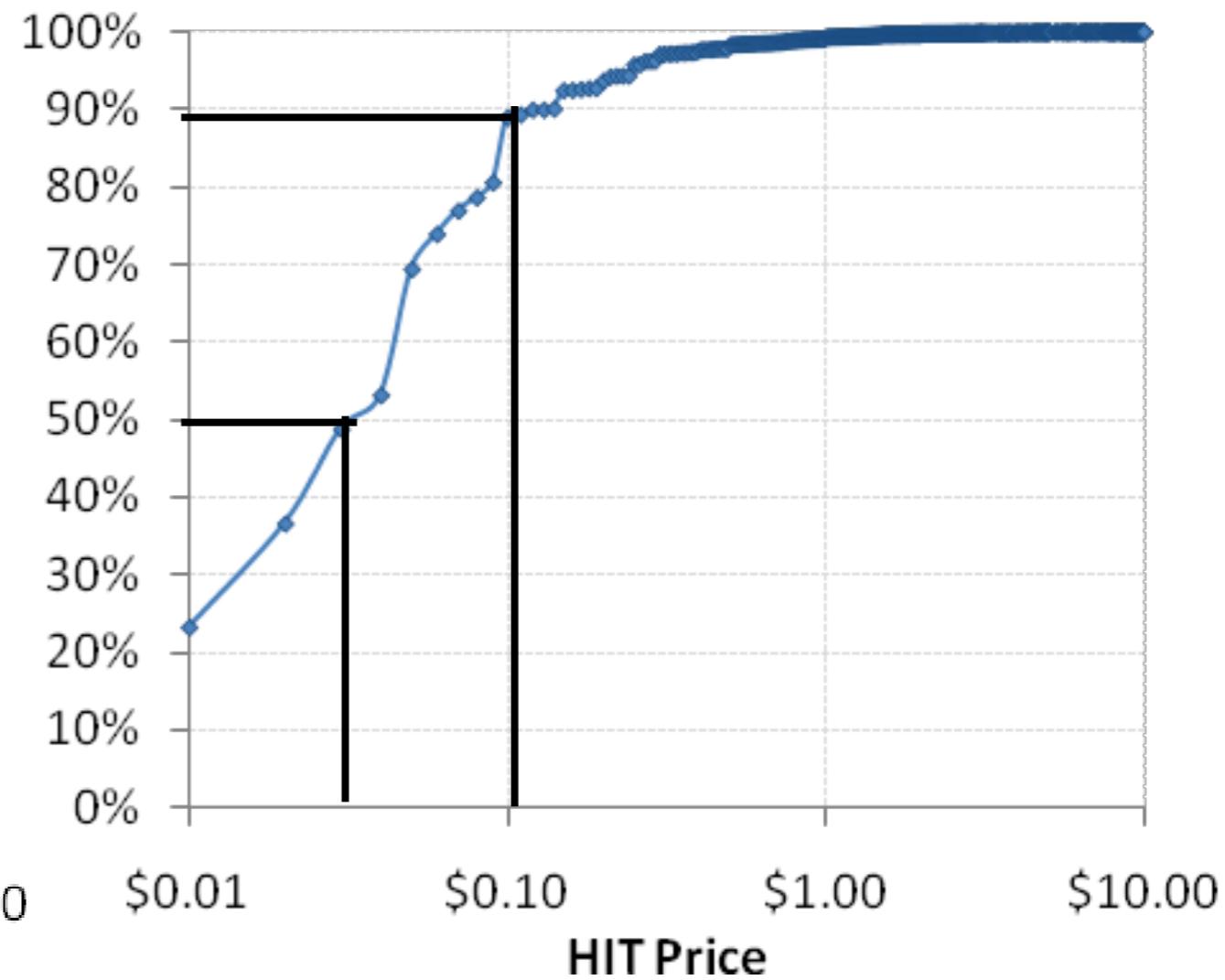
data	\$192,513	write	\$52,188	articles	\$27,841
collection	\$154,680	listings	\$48,853	category	\$26,656
easy	\$93,293	article	\$48,377	flower	\$26,131
writing	\$91,930	research	\$48,301	labs	\$26,117
transcribe	\$81,416	shopping	\$48,086	crowd	\$26,117
english	\$78,344	categorization	\$44,439	doloreslabs	\$26,117
quick	\$75,755	simple	\$43,460	crowdflower	\$26,117
product	\$66,726	fast	\$40,330	delores	\$26,117
cw	\$66,486	categorize	\$38,705	dolores	\$26,117
castingwords	\$66,111	email	\$32,989	deloreslabs	\$26,117
podcast	\$64,418	merchandise	\$32,237	entry	\$25,644
mp	\$64,162	url	\$31,819	tag	\$25,228
website	\$60,527	tagging	\$30,110	video	\$25,100
search	\$57,578	web	\$29,309	editing	\$24,791
image	\$55,013	photo	\$28,771	classify	\$24,054
builder	\$53,443	review	\$28,707	answer	\$23,856
mobmerge	\$53,431	content	\$28,319		

HITs by price

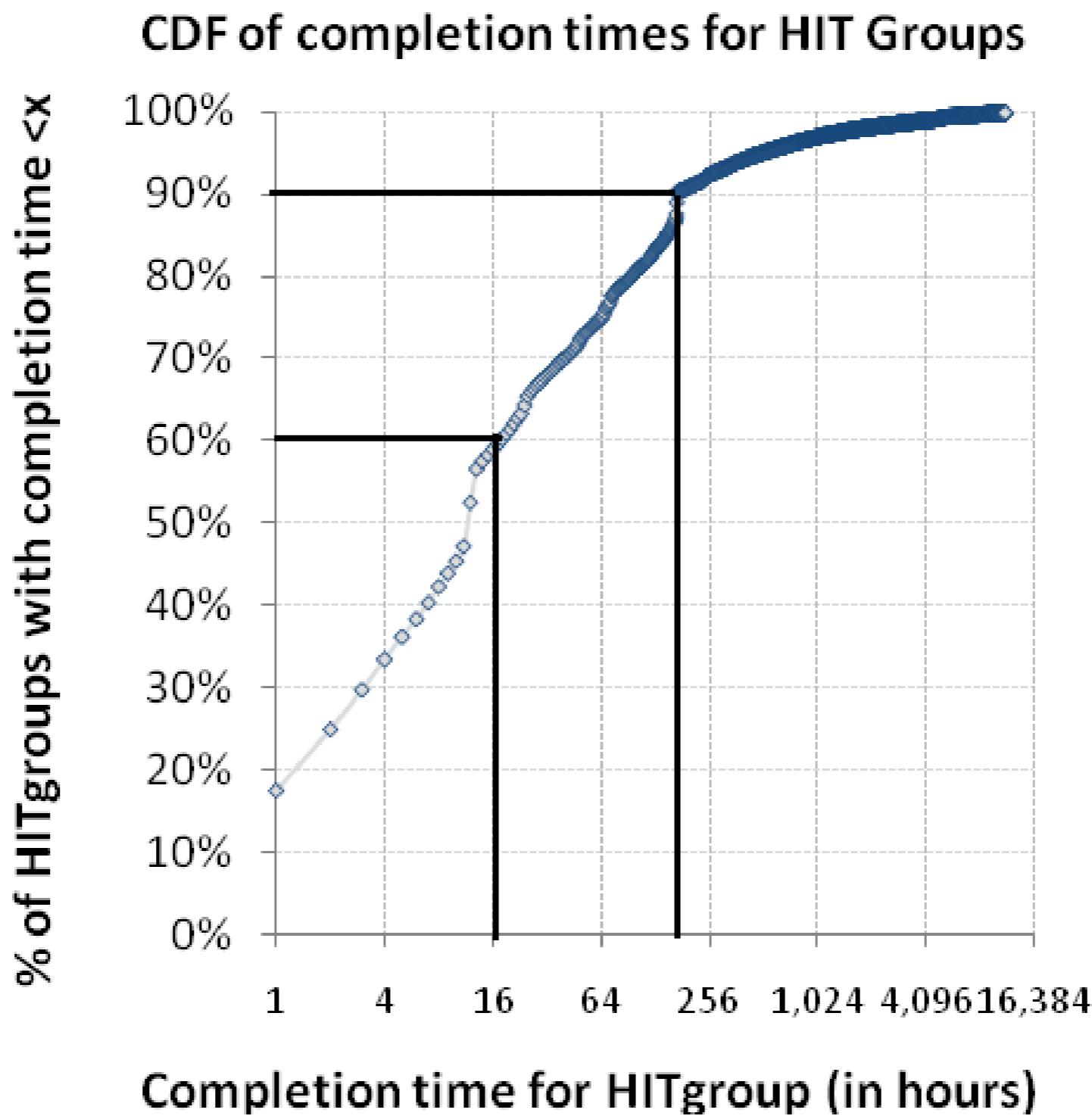
% of HITgroups vs HIT price



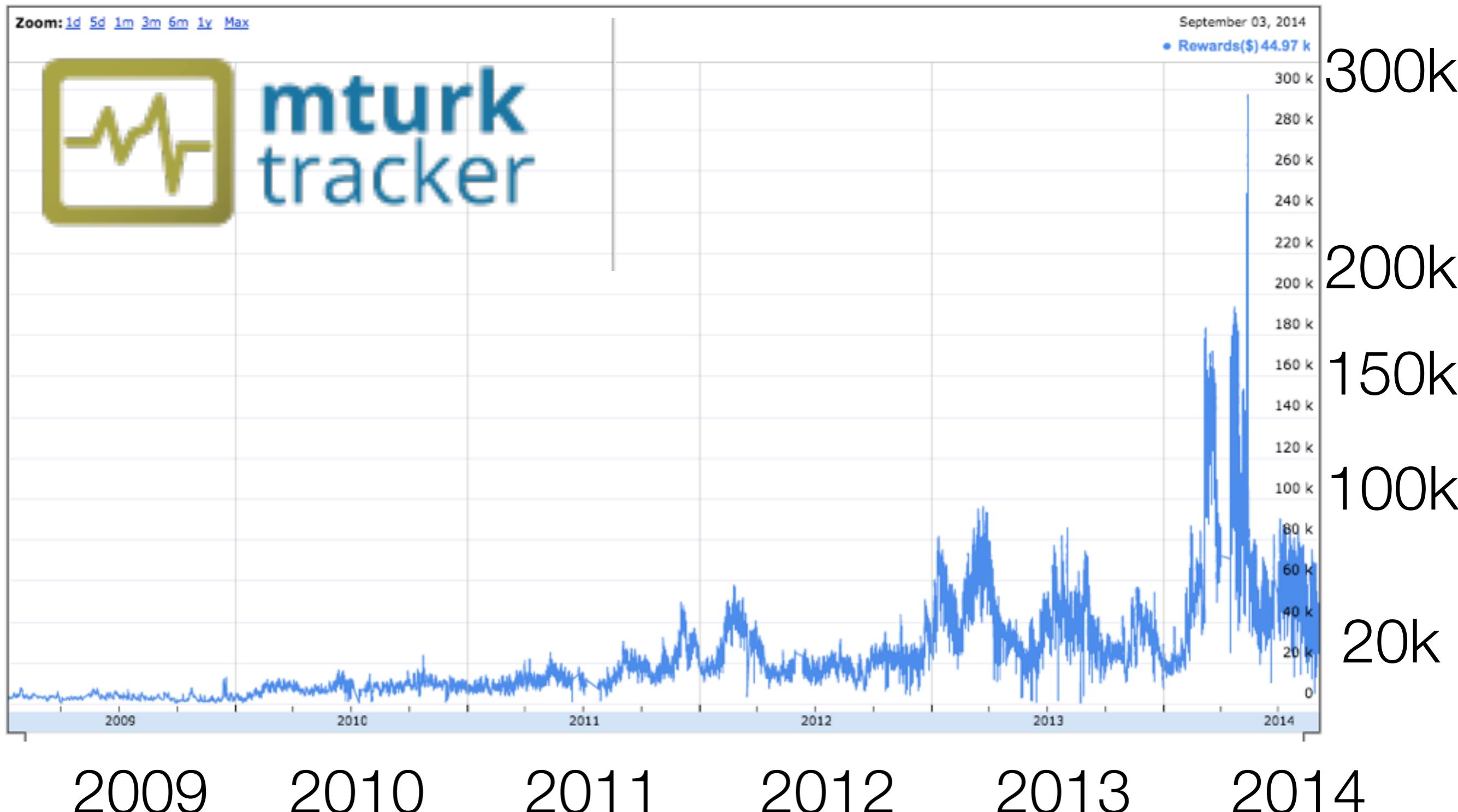
% of HITs vs HIT price



Completion times



Rewards over past 5 years



Final project ideas

- Analyze the American National Elections Survey data that I gathered before the 2012 election
- Use the data on MTurk-tracker.com to analyze the MTurk marketplace
- Help me improve [crowd-workers.com](#)

HW2: Company Profile

Reminder: Sign up for
Piazza