

Lecture 4: Knowing the Crowd Workers (Cont.)

Requesters and Tasks in Crowdsourcing Markets.

Instructor: Chien-Ju (CJ) Ho

Logistics

- Bid for Presentations
 - Make sure to bid for presentations by the end of today
- If you haven't found a team, talk to people!
- There might be minor changes to reading materials
 - Adding papers to optional readings
 - Update a few lecture titles to better reflect the included papers

Methods for Information Collection (Last Lecture)

- Direct surveys
 - Issues to keep in mind: Sampling bias; Untruthful reports
- Ethnographic analysis (as in the required reading)
- More intricate methods
 - Example 1: How to measure the communication network of the crowd
 - A web app to map networks
 - Impacts of networks to crowdsourcing
 - Example 2: How to measure the size of the crowd

Examples on Networks Induced Collaborations

- <https://blog.humancomputation.com/?p=9271>
- A half-day hackathon-style experiment

“Does the fact that workers exist in networks help them solve problems?”

- 100 workers are asked to solve TSP problems
 - Base pay: \$0.10
 - Bonus depends on the answer
 - Best answers get \$2 bonus
 - Others get bonuses \$2 – \$0.1 * deduction.
 - For every 100 mile away from the best answer, deduction++.
- Workers are explicitly told that they can collaborate

Workers started new threads to collaborate on the HIT.

MTurk Crowd

Forums

Guidelines

Members

Donate

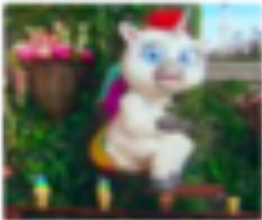
IRC Channel

Messages: 2,187

Likes: 5,561

PeachyRider, 16 minutes ago

Report



Squatty

Magical Hit Posting Unicorn

Contributor

HIT Poster

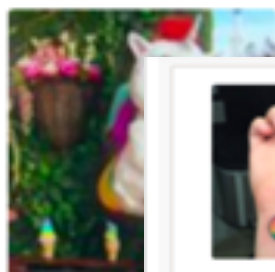
Joined: Mar 17, 2016

Messages: 6,651

Likes: 10,341

I am going to start a thread for the Suri hit, if anyone wants to join me?

Squatty, 15 minutes ago



Squatty
Magical H
Unicorn

Contri

HIT P

Joined: M

Messages



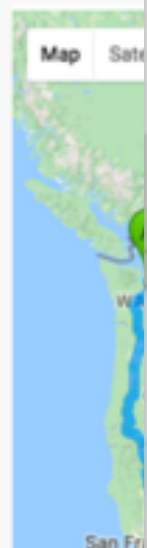
electrolyte
L'ainée de la
communauté

Administrator

Joined: Jan 10, 2016
Messages: 6,175
Likes: 12,969

Seattle to

This route



Kerek
Red Cat-Bear

Contributor

Joined: Jan 12, 2016
Messages: 7,319
Likes: 24,964



Starslip
Well-Known
Member

Contributor

Joined: Jan 19, 2016
Messages: 1,341

Jesicu said: ↑

I have Seattle, LA, Denver, Kansas City, Chicago, Atlanta, Houston, Dallas

Jesicu said: ↑

By air, the one I found is 4272.3

I think that's right. It's a straight line into Kansas City then a loop of those last five cities, and you want to go around the loop in the direction that skips the longer of the two routes out of Kansas City, which is KC->Dallas (457.9), and instead go the shorter way to Chicago (407.5).

Kerek, Sep 1, 2016 Last edited: Sep 1, 2016

#32

Kerek said: ↑

I think that's right. It's a straight line into Kansas City then a loop of those last five cities, and you want to go around the loop in the direction that skips the longer of the two routes out of Kansas City, which is KC->Dallas (457.9), and instead go the shorter way to Chicago (407.5).

I agree, I'm not sure how 4150 was arrived at with the other one as that comes out to 4488.7 for me.

Starslip, Sep 1, 2016

#

Methods for Information Collection (Today)

- Direct surveys
 - Issues to keep in mind: Sampling bias; Untruthful reports
- Ethnographic analysis (as in the required reading)
- More intricate methods
 - Example 1: How to measure the communication network of the crowd
 - A web app to map networks
 - Impacts of networks to crowdsourcing
 - Example 2: How to measure the size of the crowd

Delve into the Digital Trace

The screenshot displays the TurkNation forum homepage. At the top, there's a navigation bar with the site logo, a 'Forums' link, and various social media icons. A search bar and links for 'Help', 'Log In', and 'Register' are also present. Below the navigation bar, a banner for 'Forum: Daily mTurk HIT Threads' provides instructions for users. The main content area is a table of forum threads, categorized into 'Sticky Threads' and 'Normal Threads'. Each thread entry includes a title, a brief description, the number of replies and views, and the last post by a user. The threads cover topics like supporting the forum, survey participation, server maintenance, and advertising tips.

Title / Thread Starter	Replies / Views	Last Post By
Announcement: How you can support Turker Nation spangirl (Community Manager)		04-17-2015 Views: 54,583
Sticky Threads		
Global: When a HIT asks for your Worker ID, DON'T make these mistakes or you'll get rejected!! Hi all, I'm running a survey right now, and I'm having so many problems. But what's making me the most crazy is that some of them are totally...	Replies: 27 Views: 92,549	Juany1 03-29-2018 04:31 PM
Global: We need your help in keeping Turker Nation online The time has come to pay the server bill, and we need your help. Our Amazon Affiliate payments have dried up, so we're looking to make enough via...	Replies: 100 Views: 432,777	spangirl 08-23-2018 09:11 AM
Sticky: Tips, Tricks and Scripts to make the Daily HETs Thread easier to use Started by spangirl, 11-24-2014 09:04 PM	Replies: 91 Views: 60,601	Rosie The Robot Today, 02:17 AM
Sticky: Requester HET advertising thread Started by spangirl, 04-20-2015 06:15 PM	Replies: 0 Views: 12,766	spangirl 04-20-2015, 06:15 PM
Normal Threads		
08/21/2018 Bots Are Heartless! Tuesday Started by bromeliad, Today 06:18 AM	Replies: 278 Views: 5,405	Giskard Today, 04:02 PM

Purposes and Earnings

Why do they work?

- Money
- *“While I do find some of the HITS fun and actually learn an incredible amount by doing HITS, I do it for the cash.”*

How much they earn in a year?

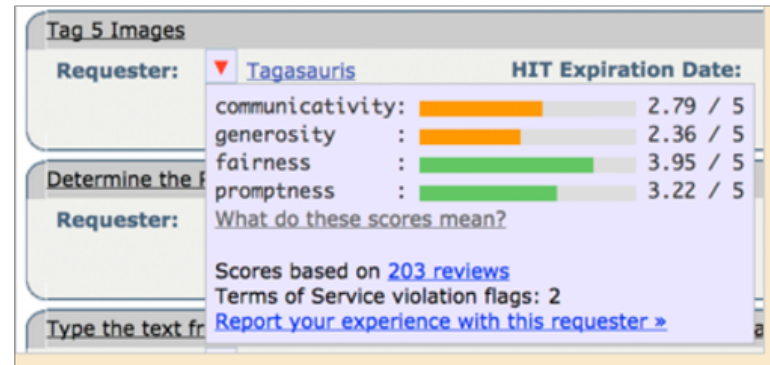
- \$1,179 to \$14,476.93. Rather low-paid!

Why do they “choose this wage”?

- From primary income to make ends meet to supplementary
- A safety-net for some!

Requester-Worker Relationship

- Workers share information on “requesters hall of fame/shame”
 - Good requesters: fair pay, quick approval, willing to communicate
 - Bad requesters: rejections without clear reasons, demeaning comments, unfair treatment
- Similar concepts lead to a plug in for “requester reputation”.
 - Turkopticon



Requester-Worker Relationship

- Workers share information on “requesters hall of fame/shame”
 - Good requesters: fair pay, quick approval, willing to communicate
 - Bad requesters: rejections without clear reasons, demeaning comments, unfair treatment
- Similar concepts lead to a plug in for “requester reputation”.
- Asymmetric power and limited communication (provided by the platform):
 - Requesters can reject work but workers can hardly appeal
 - Requesters can block workers but workers can't easily block requesters

Invisible Work

- Search for good HITs / good requesters
- Learn how to best set up computers/ tools
- Learn how to best improve skills in a certain task
- Learn how to manage turking and record keeping

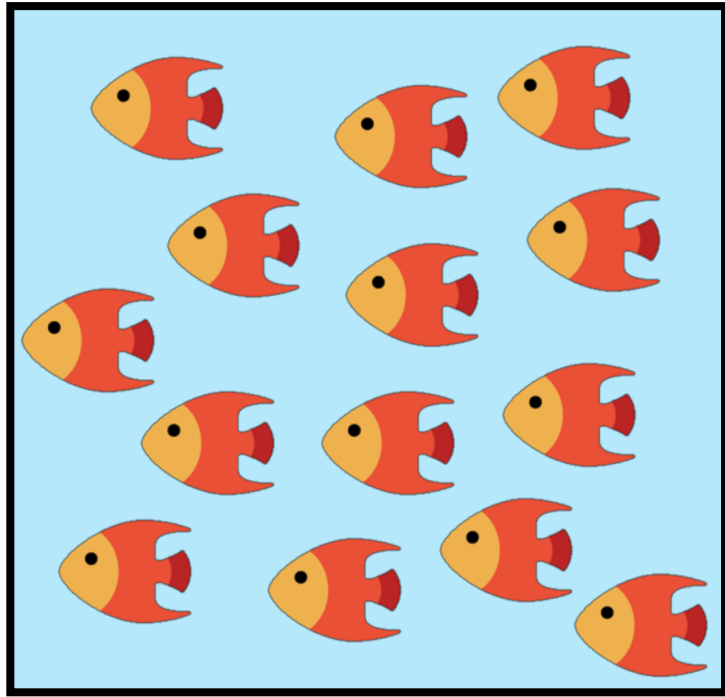
Warm-Up Discussion

- What do you think are the advantages and disadvantages of such methodology (i.e., ethnographic analysis of worker's digital trace)?
- Can you think of ways to address the limitations?

Methods for Information Collection (Today)

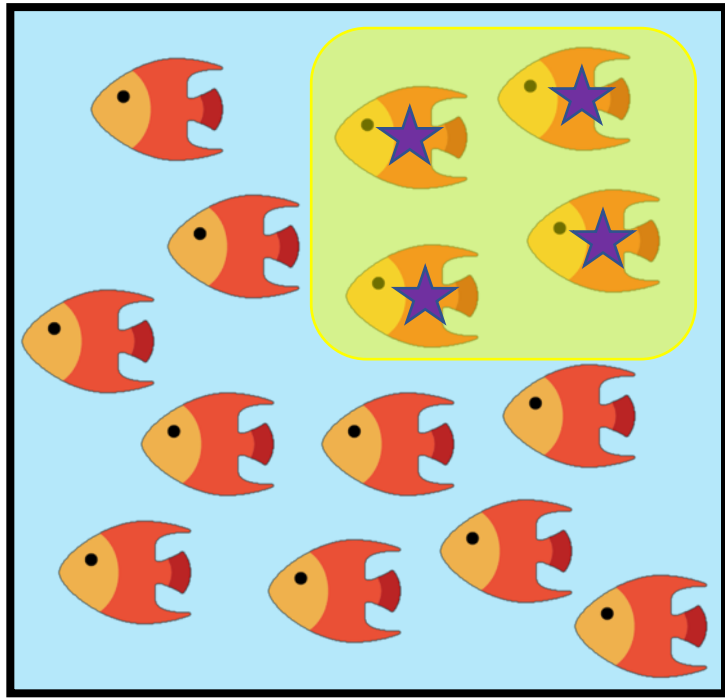
- Direct surveys
 - Issues to keep in mind: Sampling bias; Untruthful reports
- Ethnographic analysis (as in the required reading)
- More intricate methods
 - Example 1: How to measure the communication network of the crowd
 - A web app to map networks
 - Impacts of networks to crowdsourcing
 - Example 2: How to measure the size of the crowd

How Many Fishes are in the Lake?

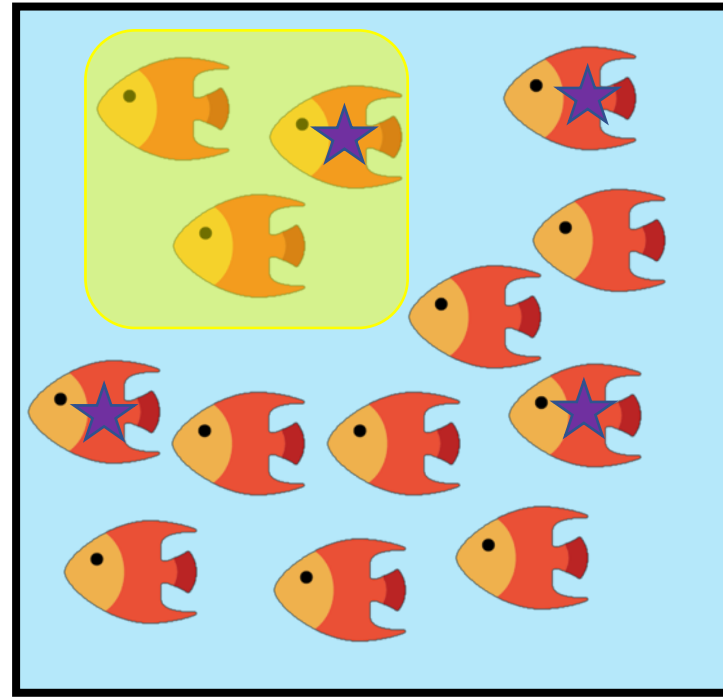


How Many Fishes are in the Lake?

The capture-recapture analysis



M fish, N_1 captured (& marked)



N_2 recaptured, n marked

$$n = \frac{N_1}{M} \times N_2$$

$$M = \frac{N_1 N_2}{n}$$

How Large is the Crowd?

- Applying the capture-recapture analysis to workers who respond to demographic surveys, we concluded that $M = 13,410$.
- What assumptions are made?
 - A1: Closed population model: No workers will leave the worker pool, and no new workers will join the pool
 - A2: Equal propensity of participation: The probability to participate in demographic survey is equal across all workers.

How Large is the Crowd?

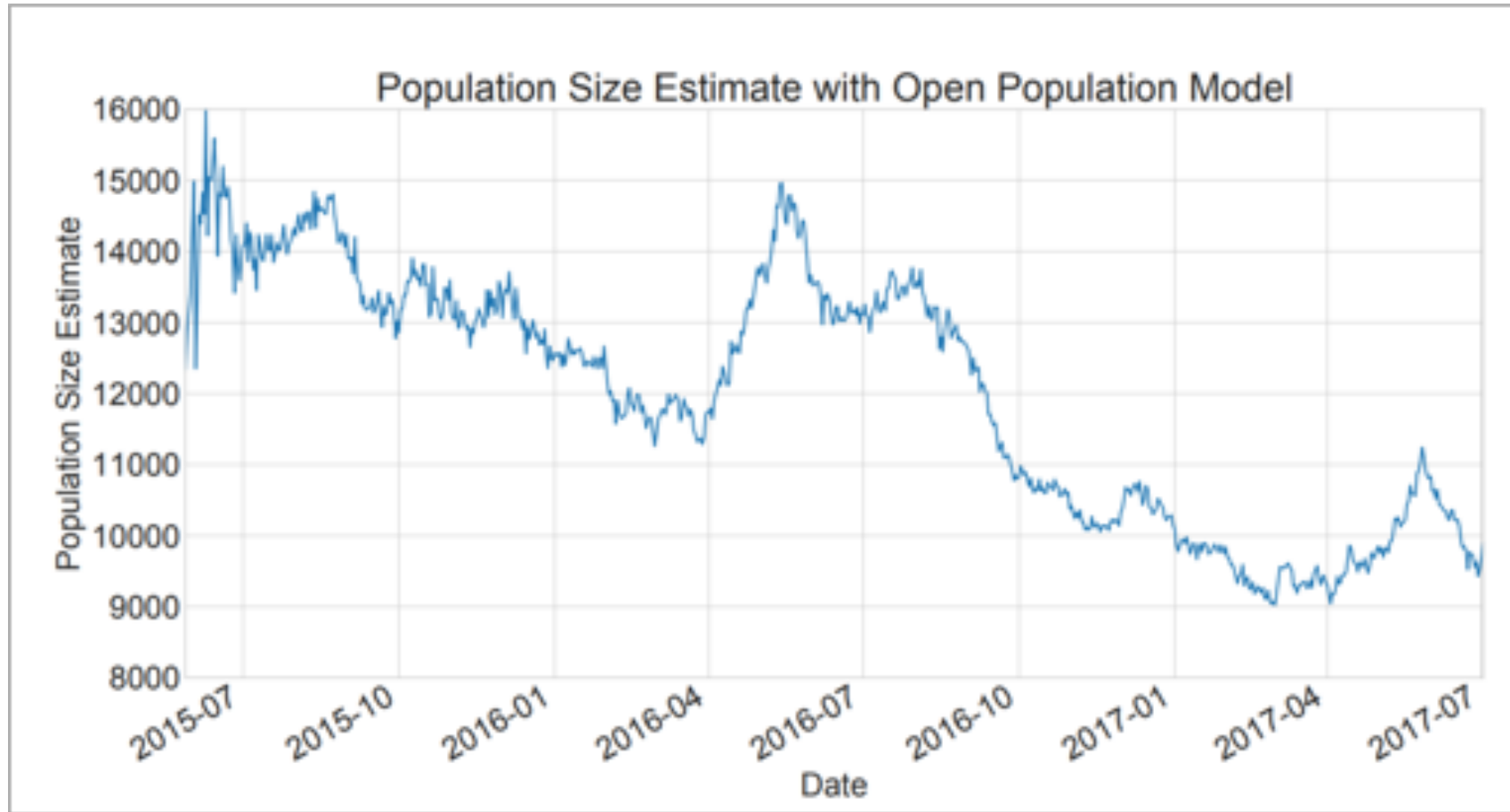
Relaxing A1: Open population model

- Survival probability: $S(t) = \exp(-\lambda t)$
- At time $d - t$, N_{d-t} workers participated in the survey
- At time d , assume that there are M_d workers in total. Among N_d workers that participated in the survey, $n_{d,t}$ workers also participated in the survey at time $d - t$.

$$n_{d,t} = \frac{N_{d-t} \exp(-\lambda t)}{M_d} \times N_d$$

$$M_d = \frac{N_{d-t} N_d \exp(-\lambda t)}{n_{d,t}}$$

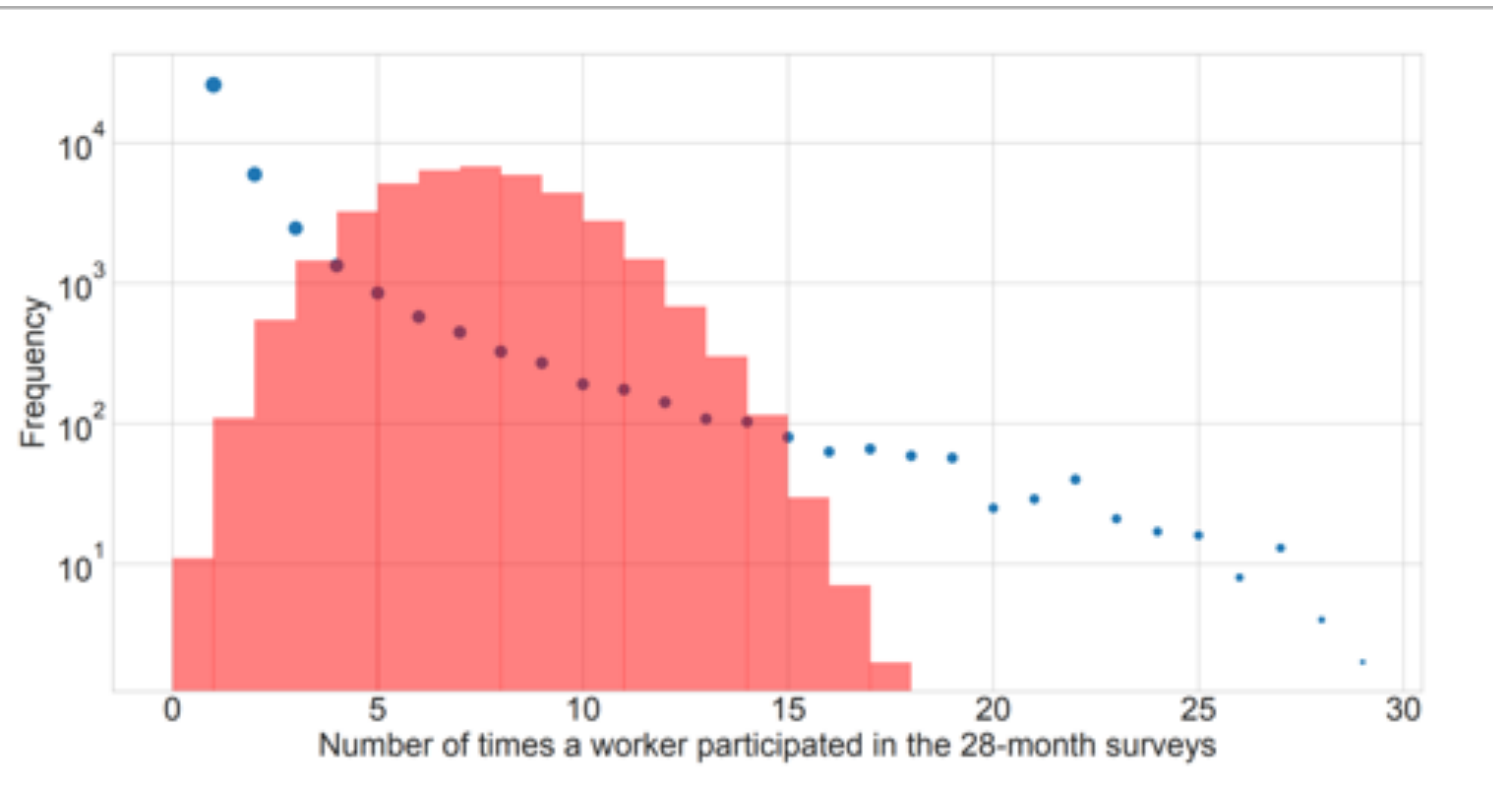
How Large is the Crowd?



Half-life of the
MTurk worker
population is 404
days

How Large is the Crowd?

- Is “A2: Equal propensity of participation” reasonable?
 - If every worker participate the survey with the same probability, we should see a binomial distribution over # participations with many surveys.



How Large is the Crowd?

Relaxing A2: Accounting for propensity of participating surveys

- Assume there are N workers in total, each worker gets a propensity score $a_i \sim p(a)$
- Capture n_1 workers, $P(\text{capture} | a_i) = 1 - (1 - \frac{a_i}{\sum_{j=1}^N a_j})^{n_1} \approx \frac{n_1 a_i}{NE[a]}$
- Recapture n_2 workers, $P(\text{capture}, \text{recapture} | a_i) = \frac{n_1 n_2 a_i^2}{N^2 E[a]^2}$
- The expected number of workers participating in both surveys:

$$m = N \int \frac{n_1 n_2 a_i^2}{N^2 E[a]^2} p(a) da = \frac{n_1 n_2}{N} (1 + \frac{\text{Var}[a]}{E[a]^2})$$

- Following this method, the estimate for the number of Mturk workers is 178,000!

Data-Driven Approaches

- Want to collect some information. Challenging to obtain enough data points to estimate the information you need.
- Propose models (i.e., making assumptions) of the process. So you only need to estimate a small number of model parameters.
- Using ML approaches to learn those parameters from data.
- Your estimate is only as accurate as the best your model can do.

Discussion

- What additional information do you want to know about Turkers?
- Can you come up with methods to collect that information?
- Does that information help us design better crowdsourcing platforms? How?

Requesters – How to Create Tasks

Creating Tasks on MTurk

- Live Demo of web interfaces
 - Setting task properties (title, description, keywords)
 - Setting payment and time
 - Setting requirements (qualifications)
 - A lot of predefined qualifications:
 - Country, approval rate, etc
 - You can set up your own qualification as well
- Setting up tasks as a webpage
 - You can use any HTML, CSS, or JavaScript to customize your layout.
 - Existing templates for common tasks are available.
- There is a [sandbox](#) website that you can play around.

Creating Tasks

- What if we want to do more complicated things that webpages can't handle?
 - Common approach:
 - Provide a link and ask workers to complete the link outside of Mturk.
 - In the outside server, give a code to Turkers when they finish the work.
 - Ask Turkers to enter the code in Mturk interface.
 - Decide on the payments based on the codes.
- There also exists API and command line tools to automate the process.

Creating Tasks

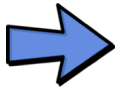
- Programming frameworks
 - [Turkit](#) (Mar 21 lecture)
 - [Automan](#)
 - ...

```
ideas = []
for (var i = 0; i < 5; i++) {
    idea = mturk.prompt(
        "What's fun to see in New York City?
        Ideas so far: " + ideas.join(", "))
    ideas.push(idea)
}

ideas.sort(function (a, b) {
    v = mturk.vote("Which is better?", [a, b])
    return v == a ? -1 : 1
})
```

Task Dynamics in MTurk

Crowd Tasks and Markets

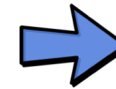


amazon **mturk** HITs Dashboard Qualifications Search All HITs

All HITs Your HITs Queue

HIT Groups (1-20 of 925) Show Blocked (0) Show Details Hide Details Items Per Page: 20

Requester	Title	HITs	Reward	Created	Actions
Chris Richmond	Search the entire website for a contact email address or form. Please also incl...	14,925	\$0.05	1d ago	Preview Qualify
James Billings	Market Research Survey	9,379	\$0.01	2m ago	Preview Accept & Work
Crowdsurf Support	Transcribe up to 35 Seconds of Media to Text - Earn up to \$0.17 per HIT!!	7,592	\$0.05	5m ago	Preview Qualify
NEL Research	Watch a 15-30 second video and share your opinions!	6,916	\$0.25	3m ago	Preview Accept & Work
Amazon Requester Inc.- Core	Indicate whether you will 1. buy one at Amazon if the other one is unavailable.	4,174	\$0.10	6d ago	Preview Qualify
Job Spotter	Collect store information of a hiring sign (WARNING: This HIT may contain adu...	3,546	\$0.05	2m ago	Preview Qualify
Vladimir	Choose correct Google Taxonomy category given product description	3,261	\$0.10	4d ago	Preview Qualify
Musicbed	Categorization Filtering for Film and Video Clips - Form 1	3,137	\$0.09	7h ago	Preview Qualify
KronoPin	Find the Website Address for a Consumer Product	2,817	\$0.05	15d ago	Preview Qualify
WatchFlower Systems	Find info from an email v1	2,556	\$0.03	1h ago	Preview Qualify
Chris Richmond	Search the entire website for a contact email address or form. Please also incl...	2,176	\$0.05	8h ago	Preview Qualify



Potential questions of interests

- What are the typical types of tasks on crowdsourcing platforms?
- How does the property of crowd tasks change over time?
- How task features influence the requester's experience?
- How task features influence the worker's experience?
- A market view: supply vs. demand

A Taxonomy of Crowd Tasks

- Surveyed 1,000 workers on CrowdFlower
- Manually annotated the type of tasks and categorized tasks into 6 classes



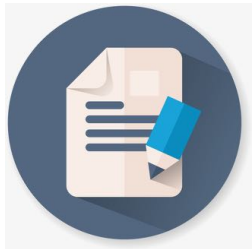
**Information
Finding**



**Verification &
Validation**



**Interpretation
& Analysis**



**Content
Creation**

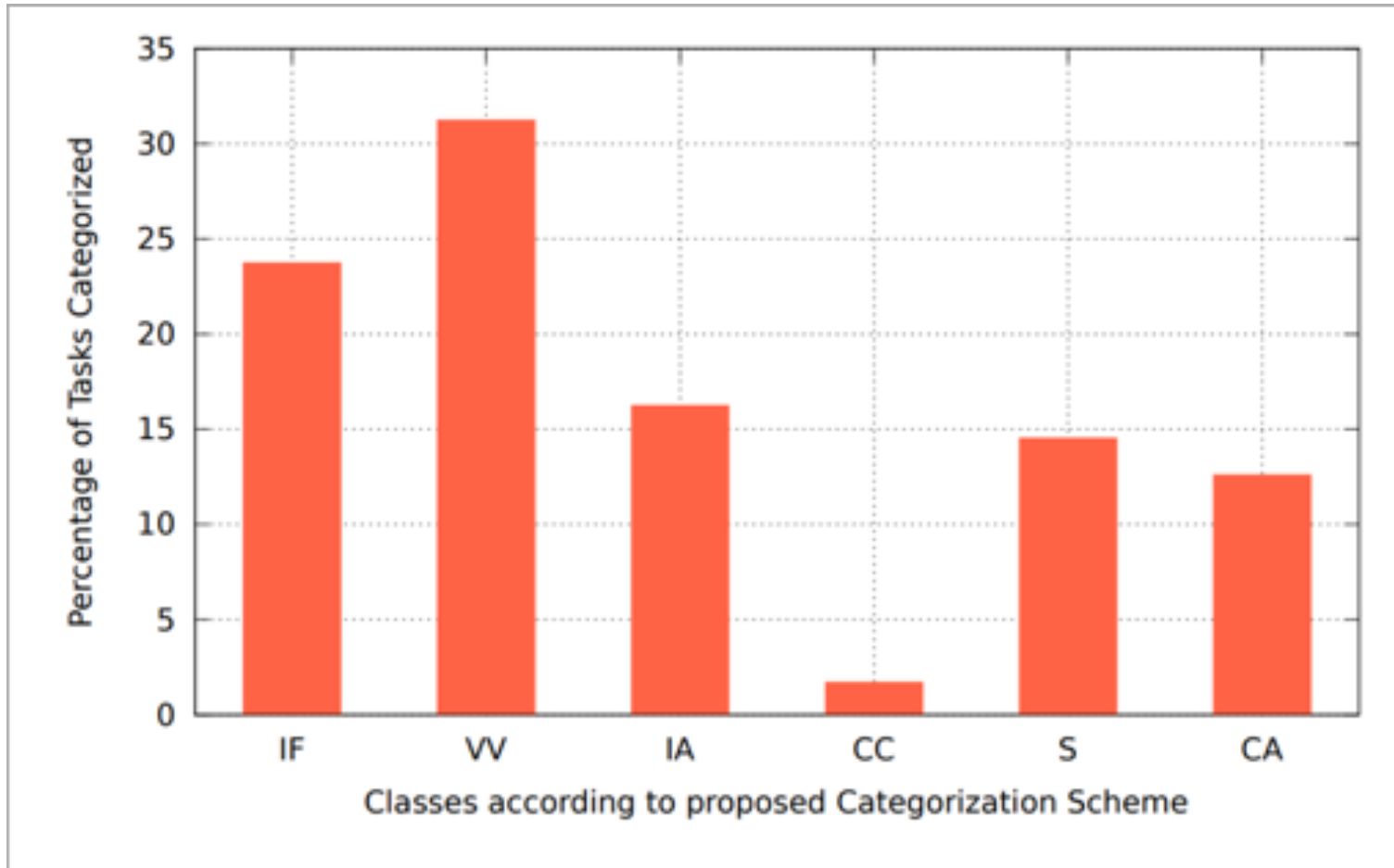


Surveys



**Content
Access**

Distribution of Task Types



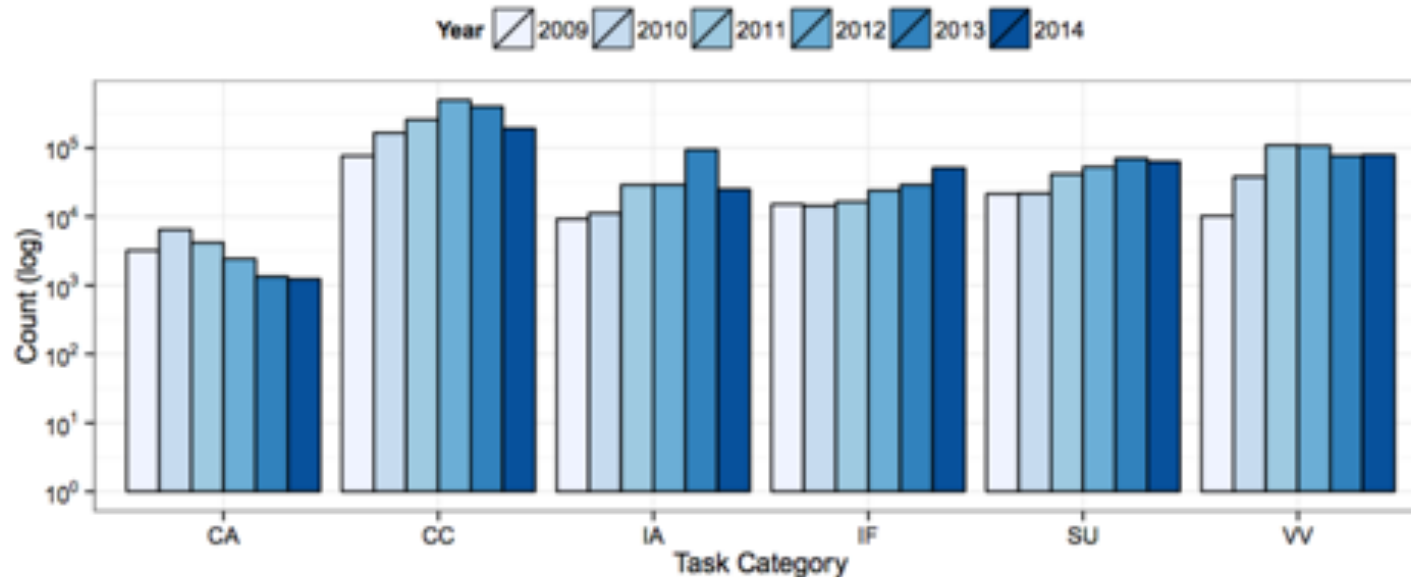
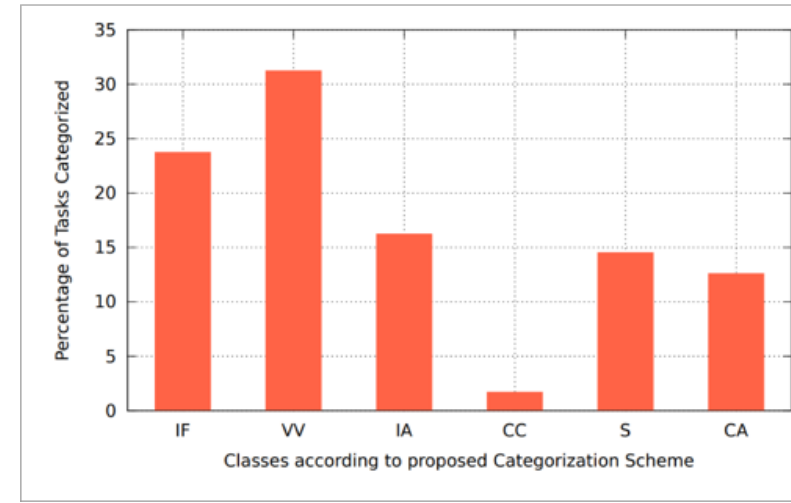
Problems?

A Case Study of MTurk

- Periodically crawled data from MTurk (data is available at MTurk-Tracker)
- Hourly aggregated data on tasks from 2009 to 2014
 - Task title
 - Task description
 - Task rewards
 - Task qualifications
 - Task keywords
 - Number of tasks available
- 2.5 million batches, 130 million tasks!

Dynamics of Task Types

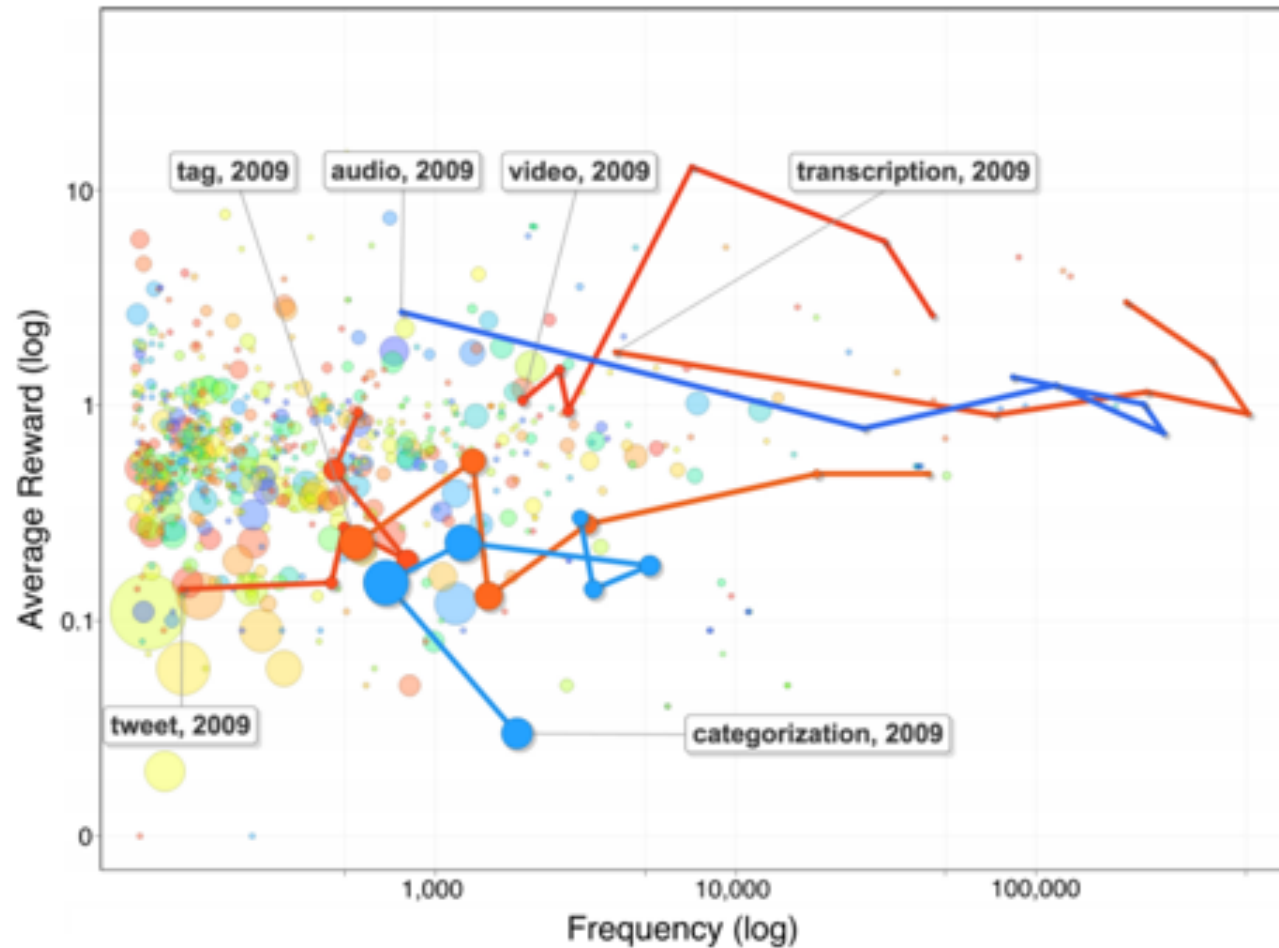
- Supervised Learning for Classifying Task Types
 - Get crowd labels for a sample of 5,000 tasks
 - Use tasks with agreement to train a SVM classifier and make predictions on all other tasks



Content Access: Decreasing
Other types: Increasing

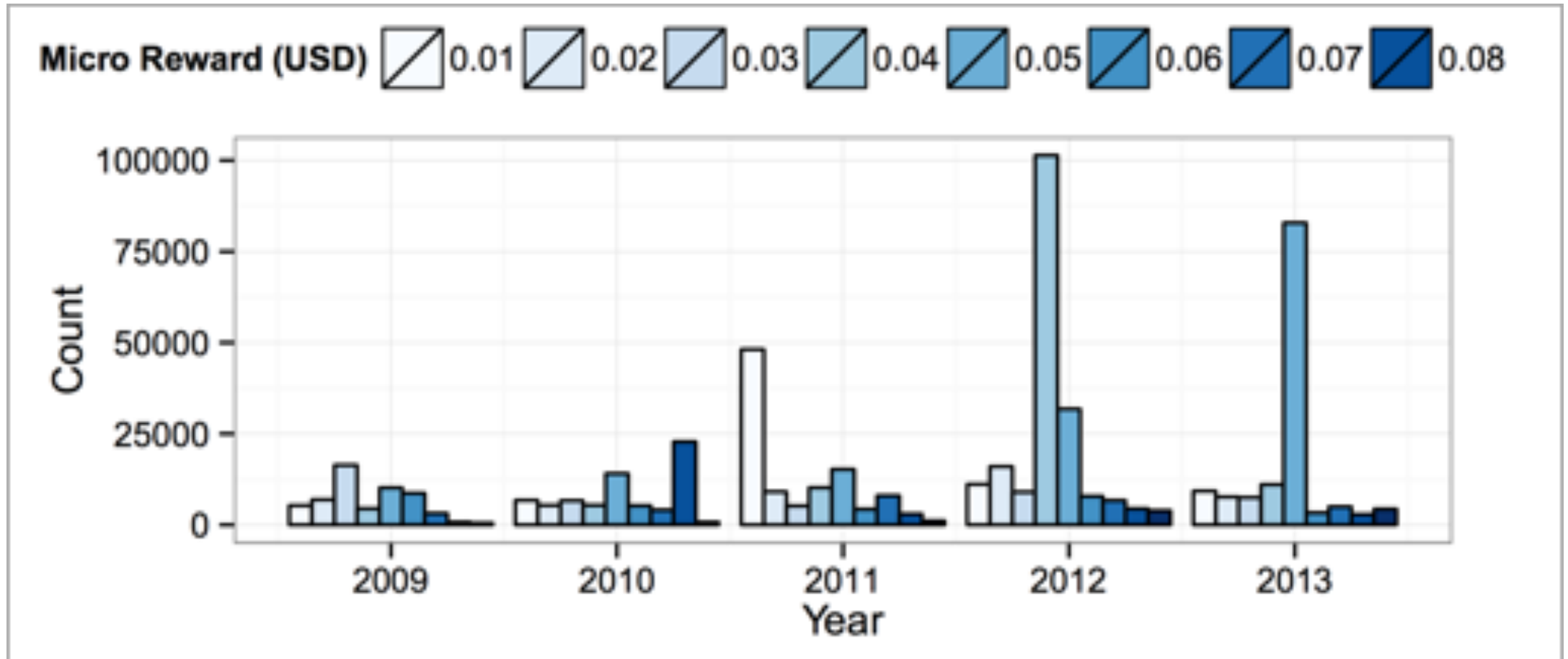
Differences between
HITs and # Assignments Per HIT

Dynamics of Task Topics



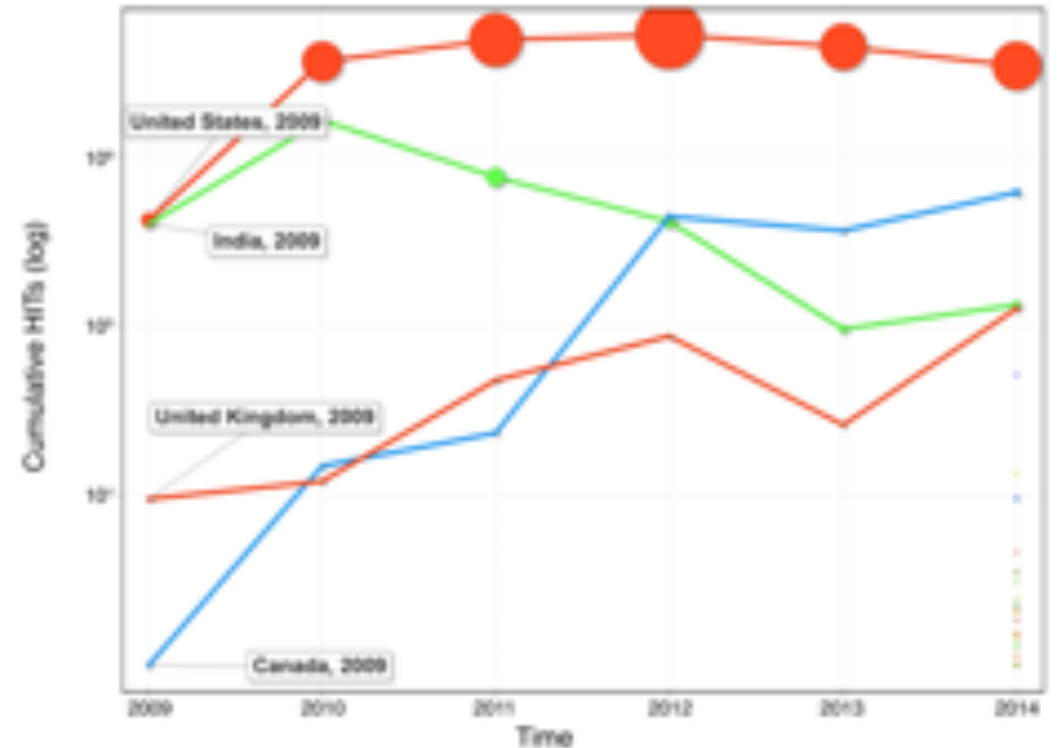
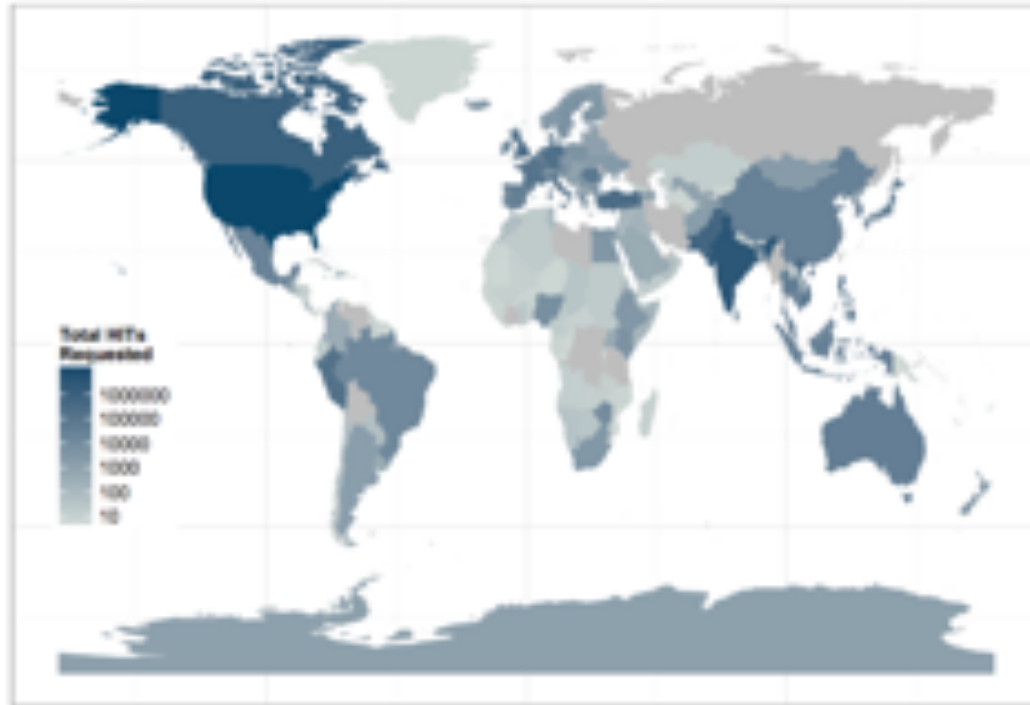
- Tasks of different topics have different "reward levels"
- Tasks related to "audio" and "transcription" increase substantially.

Dynamics of Task Rewards



Task rewards increase over time!

Dynamics of Task Qualifications (Locations)



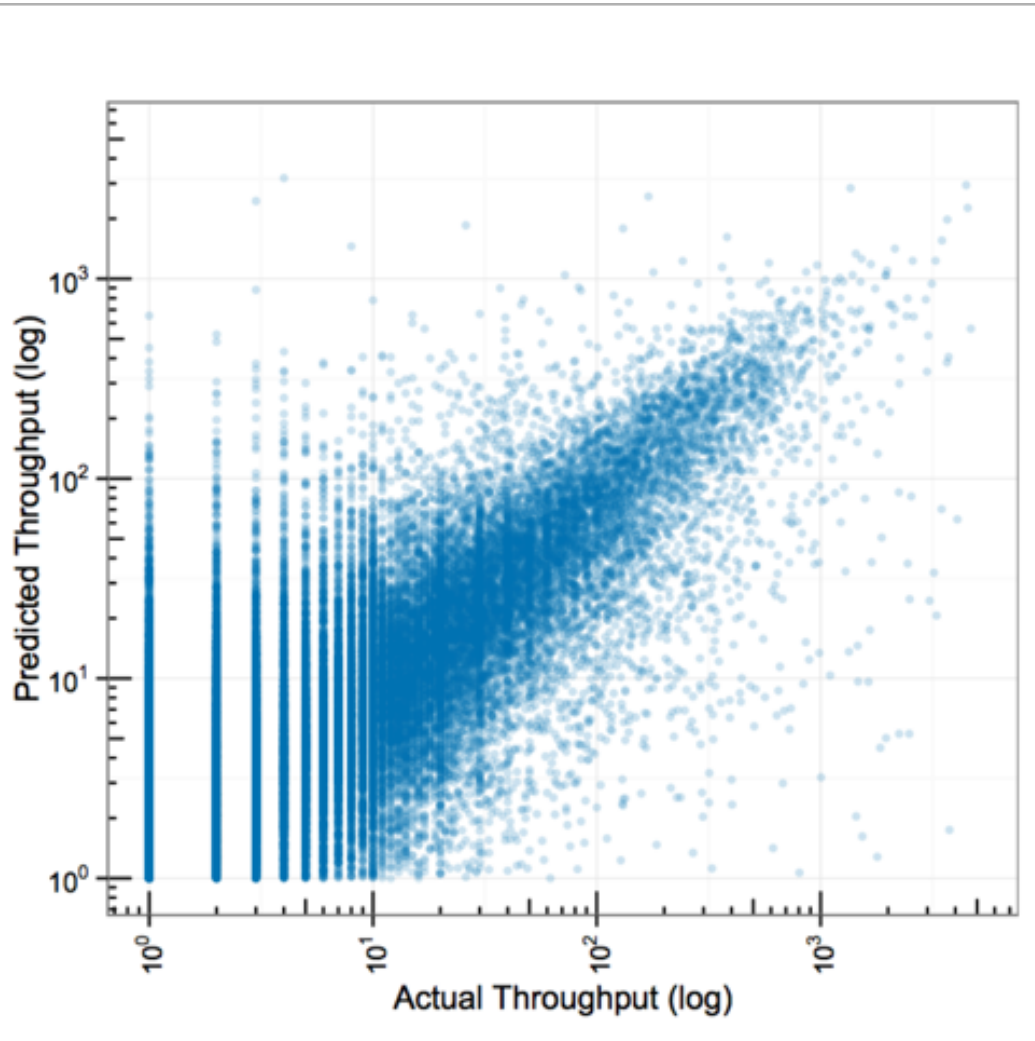
A substantial amount of tasks are exclusive to US workers, while tasks exclusive to India workers are decreasing over time!

Task Feature vs. Requester's Experience

How is the *throughout* for a batch of tasks posted by the requester influenced by task features?

Run a regression with 29 features

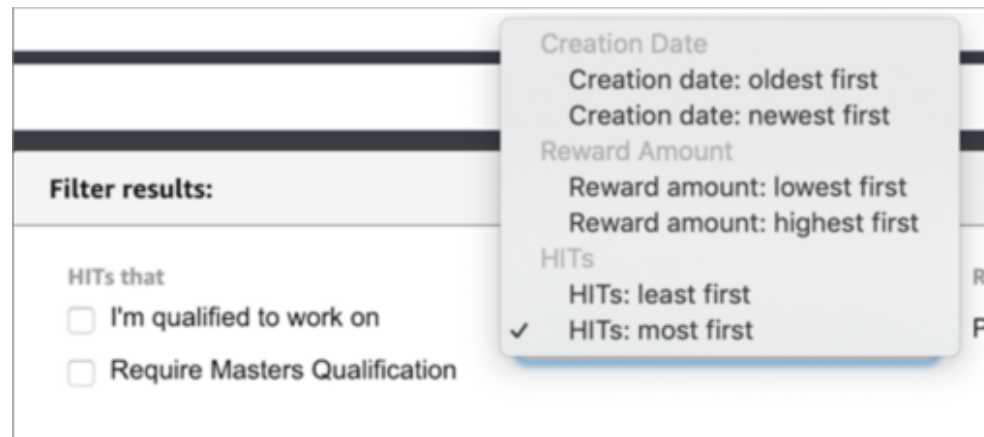
Throughput Prediction Results



- More accurate when actual throughput is large
- Important features: # of tasks available, age of the task
- Why?

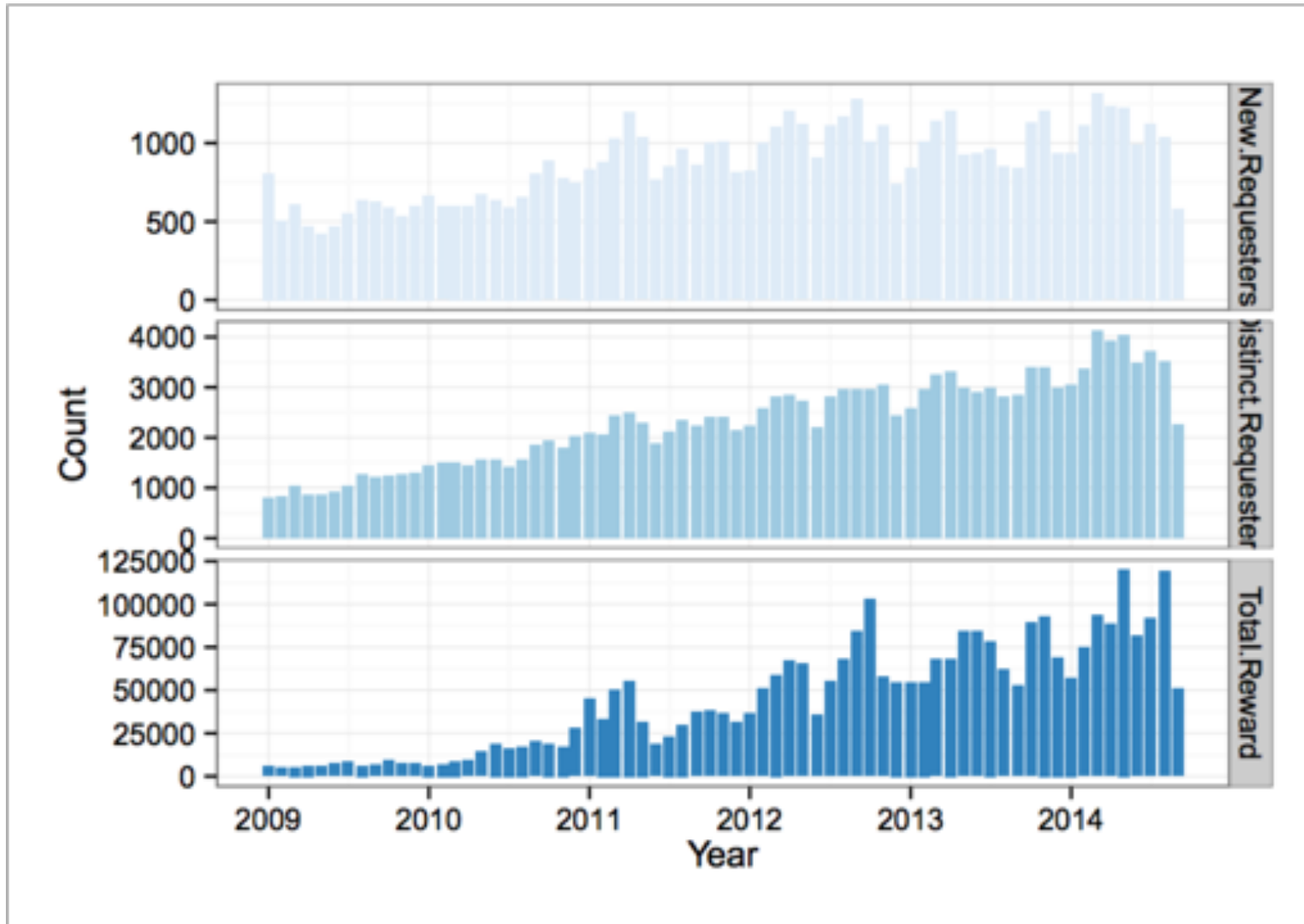
Implications of the Throughput Predictions?

- Important features: # of tasks available, age of the task
 - smaller context switch costs, etc.
 - More importantly, probably due to the MTurk interfaces in showing tasks

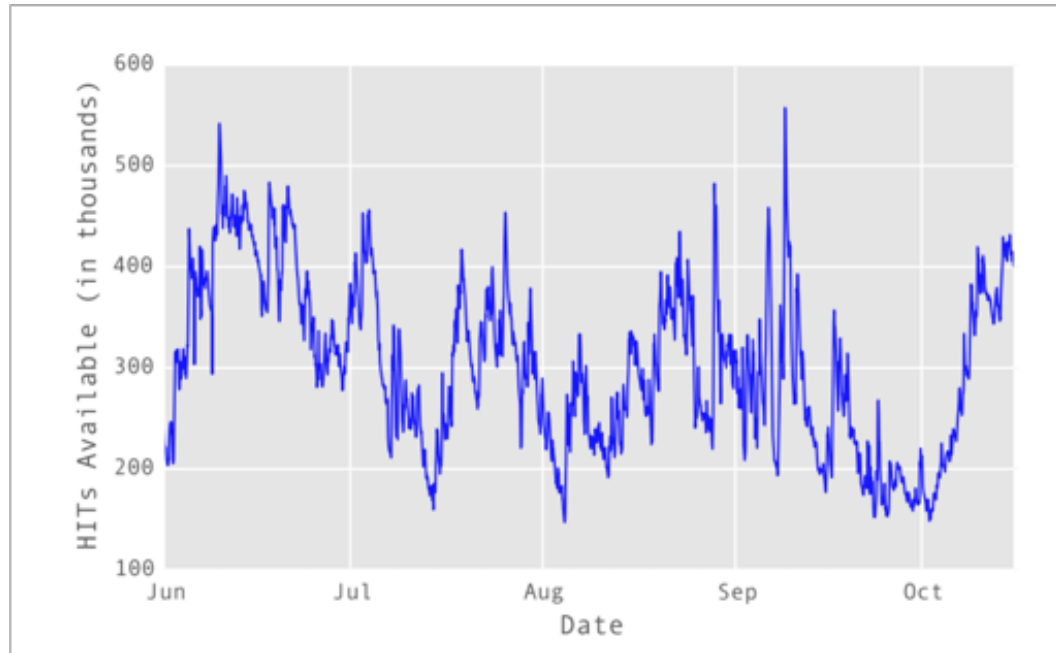


- Maybe implies the importance of task recommendation / task assignment

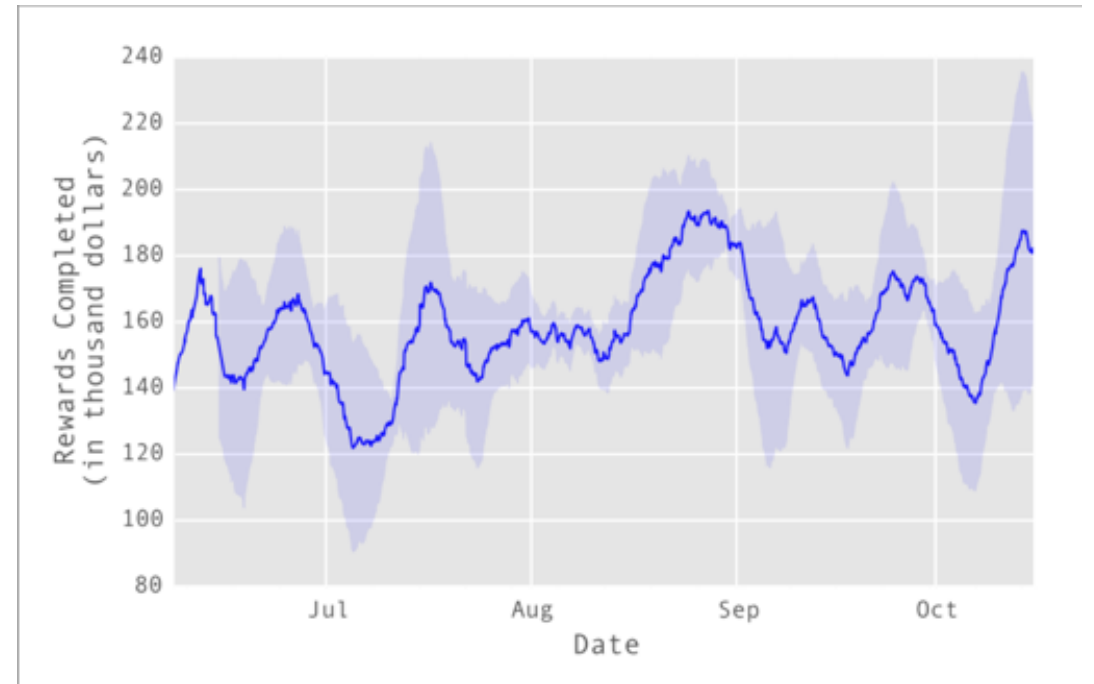
Market View: An Increasing Demand



Market View: Periodicity in Supply and Demand

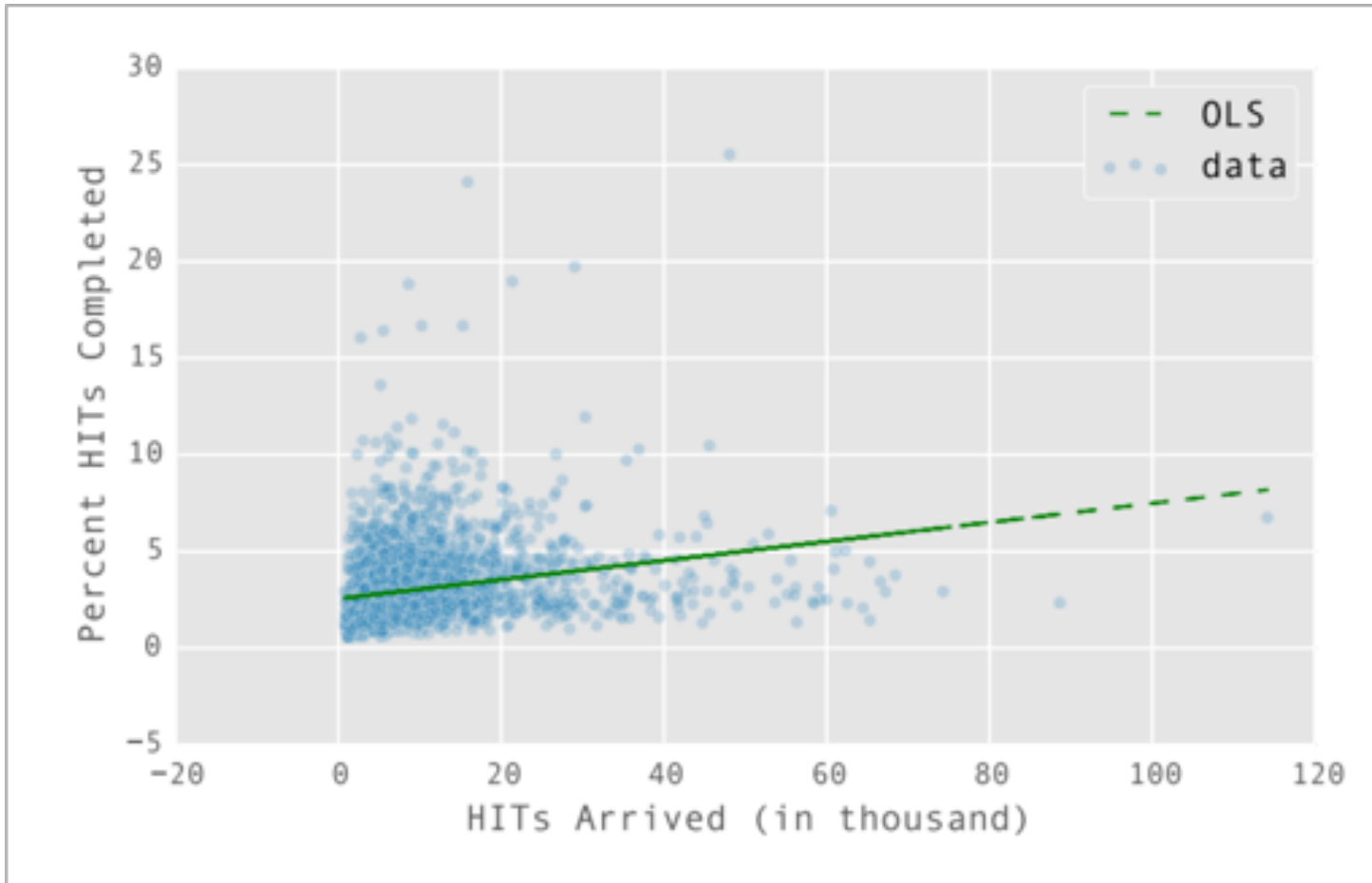


Demand



Supply

Market View: Supply Increases with Demand



- When the demand in an hour increases by 10K, the percentage of completed tasks increase by 0.5%.

Discussion

- If you had this data, what other analyses you want to conduct to better understand crowdsourcing tasks and markets?
- Do you need to collect some additional data for these analyses? If yes, what data do you need?

Beyond Microtask Platforms

- Fiverr: <https://www.fiverr.com/>
- Upwork: <https://www.upwork.com/>
- What are the questions you want to ask for understanding the dynamics of these crowdsourcing platforms (especially questions that are different from what you will ask for microtask platforms)?
- What data you would need to collect to answer these questions?