



PROGRAMMING FOR SOCIAL SCIENTISTS WITH INTERPOLL

Ben Livshits
Todd Mytkowicz
Microsoft Research

InterPoll 101

We want to give access to polls and survey to the **regular developer**.

```
1 var people = new MTurkQueryable<Person>(true, 5, 100, 2);
2 var liberalArtsPairs = from person in people
3   where person.Employment == Employment.STUDENT
4   select new {
5     Person = person,
6     Value = person.PoseQuestion<bool>(
7       "Are you a liberal arts major?")
8   };
```

This is a LINQ query for asking **college students** whether they are **liberal arts majors**

We want to make access to **human-generated data** as easy as access to **databases**.

SIMPLE LINQ QUERIES

```
var femaleHeight = from person in people where person.Gender ==  
Gender.FEMALE select person.PoseQuestion<int>("What is your height?");  
  
var maleHeight = from person in people where person.Gender ==  
Gender.MALE select person.PoseQuestion<int>("What is your height?");
```

```
if (maleHeight.ToRandomVariable() > femaleHeight.ToRandomVariable()) {  
    Console.WriteLine(  
        "Males are taller than females, according to a t-test.");  
}
```

C# + LINQ

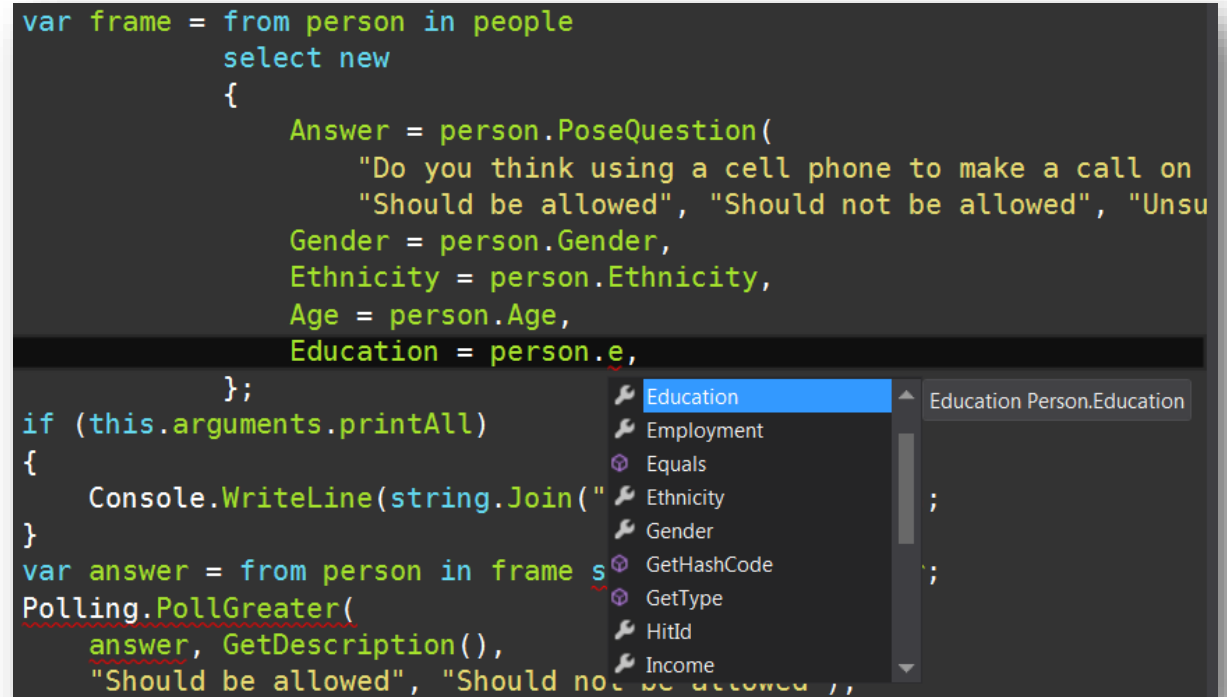
Developers can continue using familiar tools such as Visual Studio, with powerful IntelliSense support and type checking, to avoid making mistakes.

LINQ (Language-integrated queries) is the only way to create surveys, but it has several important advantages.

In InterPoll, we are thinking a lot about the **developer experience**. LINQ is familiar to developers and does not require learning a new domain-specific language.

LINQ queries integrate well with the rest of the application.

```
var frame = from person in people
            select new
            {
                Answer = person.PoseQuestion(
                    "Do you think using a cell phone to make a call on
                    "Should be allowed", "Should not be allowed", "Unsu
                Gender = person.Gender,
                Ethnicity = person.Ethnicity,
                Age = person.Age,
                Education = person.e,
            };
if (this.arguments.printAll)
{
    Console.WriteLine(string.Join("
}
var answer = from person in frame s
Polling.PollGreater(
    answer, GetDescription(),
    "Should be allowed", "Should not be allowed",
```

A screenshot of a code editor showing C# LINQ code. The code defines a query 'frame' that selects new objects with properties Answer, Gender, Ethnicity, Age, and Education. Below the query, there is an if-statement for printing all arguments and a call to Polling.PollGreater. An IntelliSense dropdown menu is open over the 'Education' property assignment, showing a list of members including Education, Employment, Equals, Ethnicity, Gender, GetHashCode, GetType, HitId, and Income. The 'Education' member is highlighted in blue, and a tooltip for 'Education Person.Education' is visible on the right side of the dropdown.

DEMO

```
1 using ...
4
5 namespace Microsoft.Research.RiSE.InterPoll
6 {
7     public partial class Runner
8     {
9         [TestMethod]
10        public void EmploymentSurvey() {
11
12        }
13    }
14 }
```

SAMPLE SURVEYS

Survey

- What is currently preventing
- Time Available
 - Desire and Motivation
 - Weather
 - Laws of Thermodynam
 - Physical Disability

Survey

- What has happened to the US
- Smaller
 - Larger
 - Same as it was a year

Survey

Do you shop locally?

- Always
- Never
- Once in a while
- Usually
- About half the time

Do you make at least one purchase a day at chain stores?

- Yes
- No

Do you shop at local stores daily?

- Yes
- No

Do you consider yourself to be a supporter of small business?

- Yes
- No

Survey

I feel tense or "wound up"

- Most of the time
- A lot of the time
- From time to time, occasionally
- Not at all

I still enjoy the things I used to enjoy

- Definitely as much
- Not quite so much
- Only a little
- Hardly at all

I get a sort of frightened feeling as if something awful is about to happen

- Very definitely and quite badly
- Yes, but not too badly
- A little, but it doesn't worry me
- Not at all

I can laugh and see the funny side of things

- As much as I always could
- Not quite so much now
- Definitely not so much now
- Not at all

Worrying thoughts go through my mind

- A great deal of the time
- A lot of the time
- From time to time but not too often
- Only occasionally

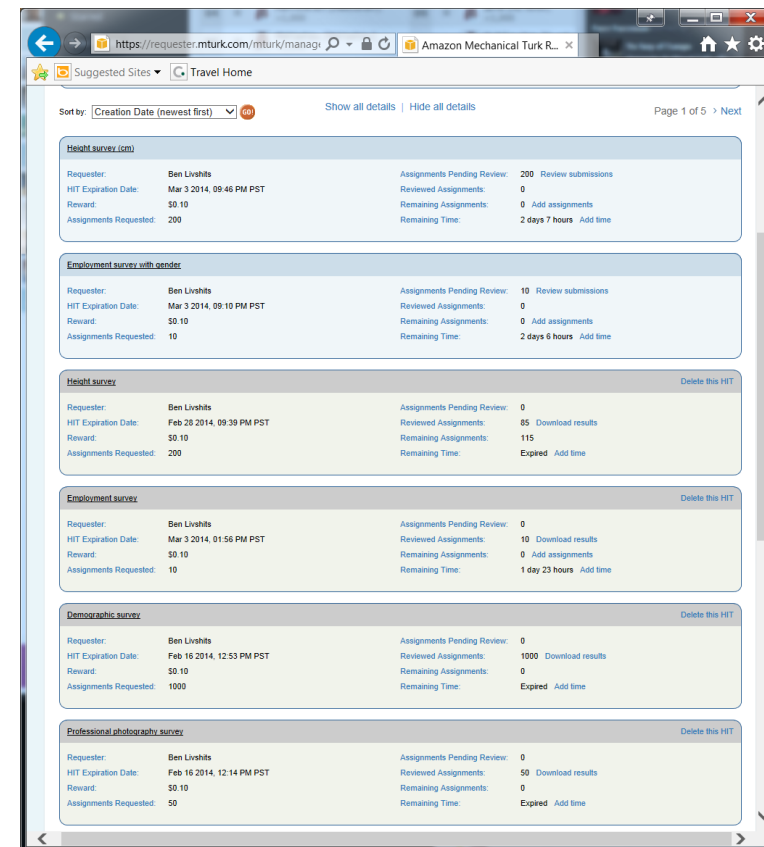
I feel cheerful

- Not at all
- Not often
- Sometimes

MECHANICAL TURK BACK-END

InterPoll uses Mechanical Turk as a back-end

Others like UHRS are possible



The screenshot displays the Amazon Mechanical Turk back-end interface. The browser address bar shows the URL <https://requester.mturk.com/mturk/manage>. The page is titled "Amazon Mechanical Turk R..." and includes navigation options like "Suggested Sites" and "Travel Home". The main content area shows a list of HITs (Human Intelligence Tasks) with the following details:

HIT Title	Requester	HIT Expiration Date	Reward	Assignments Requested	Assignments Pending Review	Reviewed Assignments	Remaining Assignments	Remaining Time	Additional Info
Height survey (cm)	Ben Livshits	Mar 3 2014, 09:46 PM PST	\$0.10	200	200	0	0	2 days 7 hours	Review submissions, Add time
Employment survey with gender	Ben Livshits	Mar 3 2014, 09:10 PM PST	\$0.10	10	10	0	0	2 days 6 hours	Review submissions, Add time
Height survey	Ben Livshits	Feb 28 2014, 09:39 PM PST	\$0.10	200	0	85	115	Expired	Download results, Add time
Employment survey	Ben Livshits	Mar 3 2014, 01:56 PM PST	\$0.10	10	0	10	0	1 day 23 hours	Download results, Add time
Demographic survey	Ben Livshits	Feb 16 2014, 12:53 PM PST	\$0.10	1000	0	1000	0	Expired	Download results, Add time
Professional photography survey	Ben Livshits	Feb 16 2014, 12:14 PM PST	\$0.10	50	0	50	0	Expired	Download results, Add time

IS THIS NOT A SOLVED PROBLEM?

SurveyMonkey

Create surveys
anywhere, anytime.

4:21 PM 100%

Edit

Add Logo

Customer Feedback Survey

1. How well did the customer service agent answer your questions?

Extremely well

Very well

Moderately well

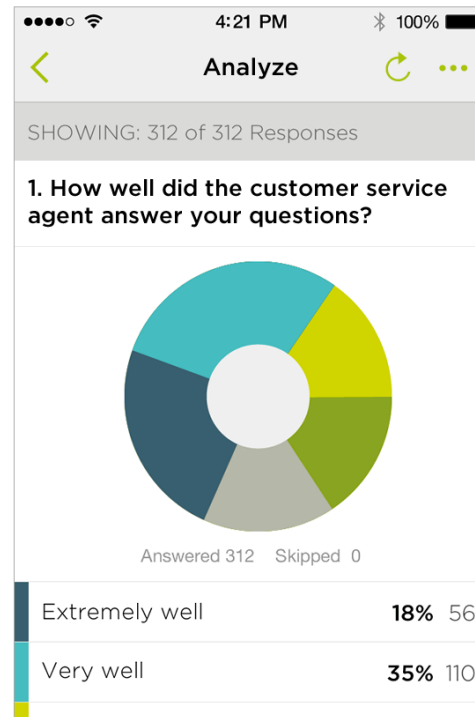
Slightly well

Not at all well

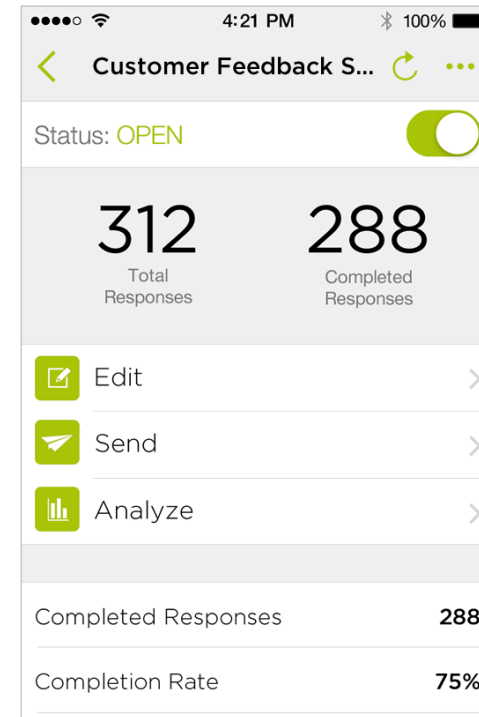
+ Add Question

+ Add Page

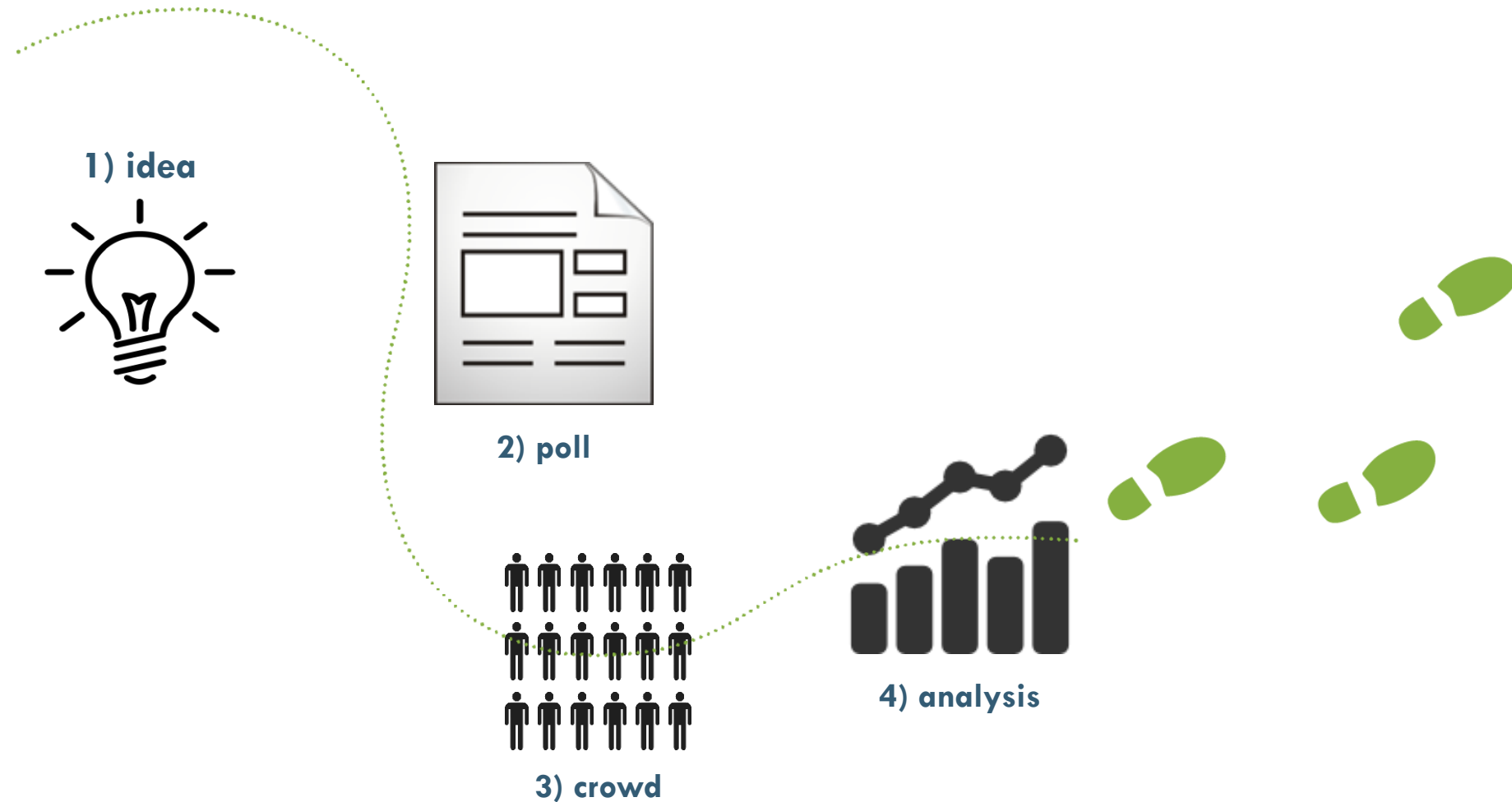
Analyze your survey
results on the go.



Track your results
in real-time.



OUR FOCUS IS ON AN END-TO-END PROCESS



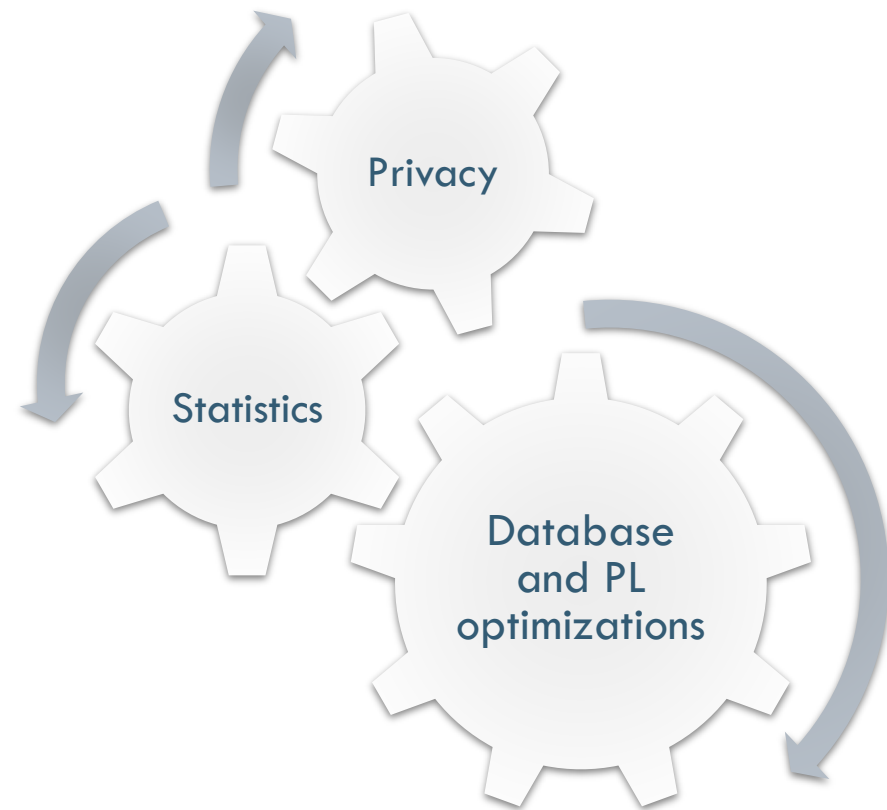
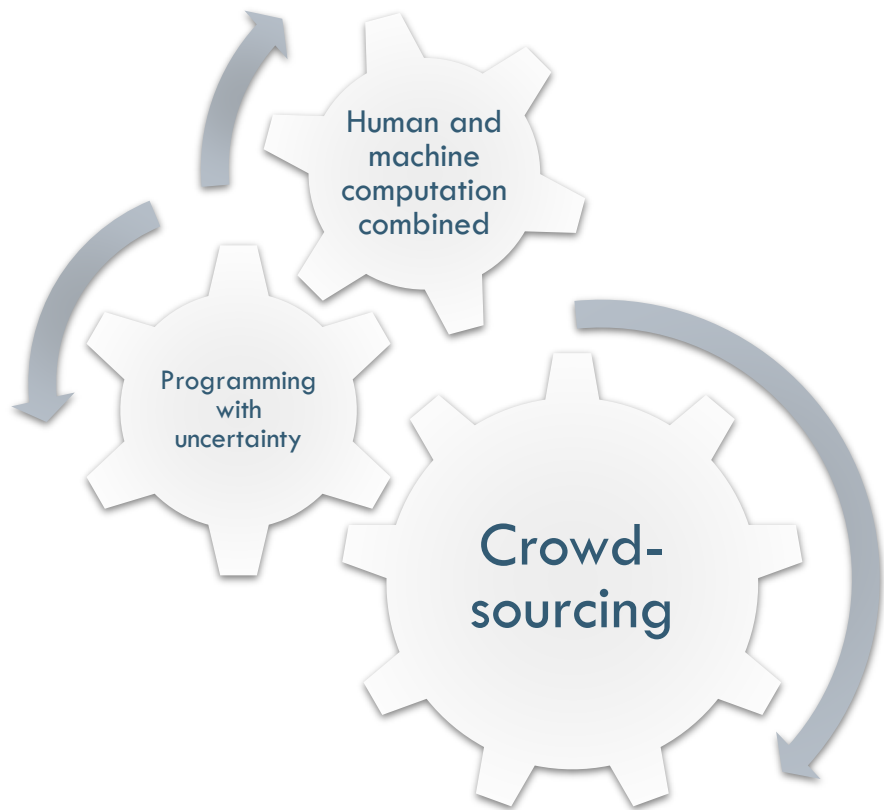
KEY DIFFERENCES BETWEEN THEM AND INTERPOLL

Lack of programmability

InterPoll is considerably cheaper

We give statistical significance to the results through **unbiasing** and **power analysis**

A SMORGASBORD OF TOPICS





One of the **wonderful** things about InterPoll is that it brings together research fields that are **not** frequently combined.

This requires researchers who are **not afraid** to step outside their comfort zone.



“Experimental psychology is
the **study** of the **college sophomore**”

Quinn McNemar, 1946



Traditional polling is a painstaking and **time-consuming** process.

Not everything can be assessed through computer-based polling, but many things can be.

Recent studies show that online polling is frequently better than RDD polls and are good at avoiding the **interviewer bias**.



Recent results suggest that online polling is actually **cheaper**, **faster**, and is devoid of some of the **biases** such as the interviewer bias present in traditional face-to-face or telephone surveys.

Moreover, in recent years, traditional random-dialing surveys suffer from the fact that **young people** often do not have a land line.



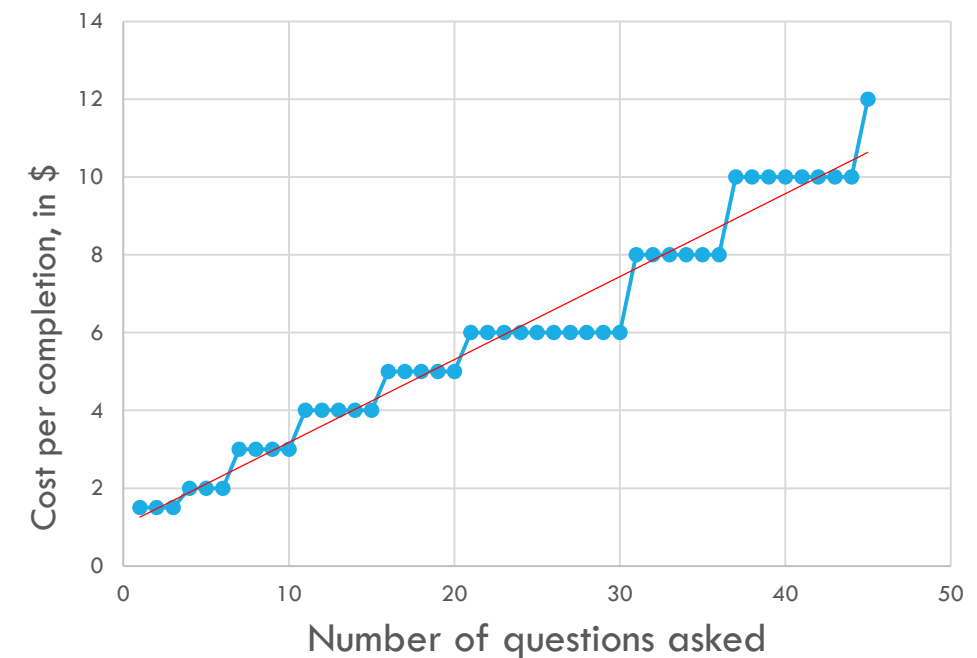
COSTLY... ESPECIALLY AT SCALE

SurveyMonkey Audience pricing

A technology accessory company wants feedback on their latest iPhone case design. They have a 14-question survey and would like 200 responses from an audience of only female iPhone users, delivered on the 2-business day schedule.

14 question survey	\$1.50 per response
2 specific targeting options added	\$1.25 per response (Gender targeting & iPhone ownership targeting)
2-business day turnaround	\$1.00 per response
Project cost for 200 responses	\$750.00 (\$3.75 per response)

Instant.ly cost per completed survey



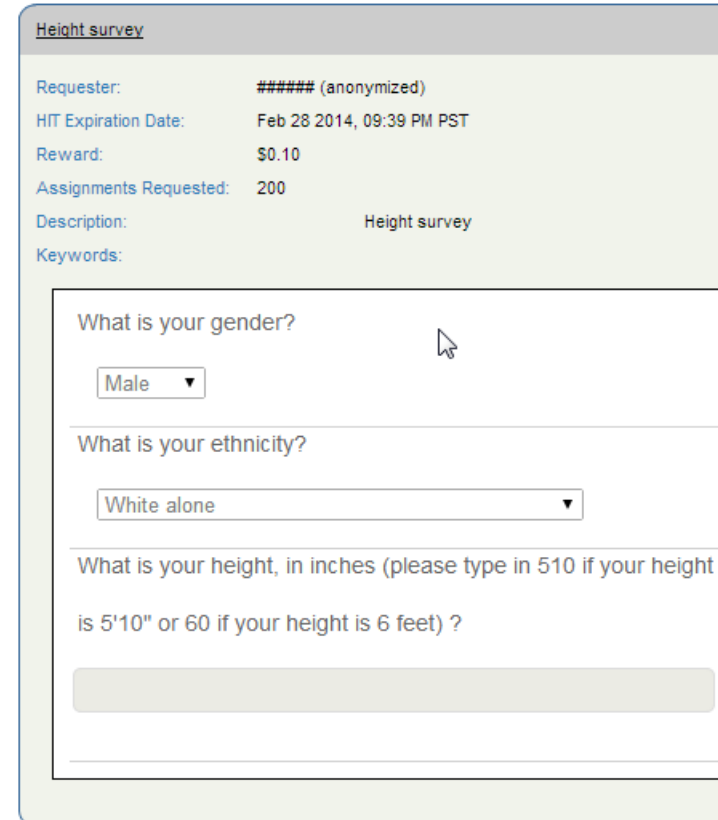
MECHANICAL TURK (MTURK)

Our main backend is Mechanical Turk

Generally can get answers to most polls for \$.10/survey completion

Latency drops as a result of higher rewards, but the quality generally doesn't seem to increase

InterPoll takes care of creating appropriate HITs, generating XML forms for the UI that appears on Mechanical Turk and collecting the results



The screenshot shows a Mechanical Turk HIT interface for a survey titled "Height survey". The interface includes a header with the title and a list of details: Requester (##### (anonymized)), HIT Expiration Date (Feb 28 2014, 09:39 PM PST), Reward (\$0.10), Assignments Requested (200), Description (Height survey), and Keywords. Below the details is a form with three questions: "What is your gender?" with a dropdown menu showing "Male", "What is your ethnicity?" with a dropdown menu showing "White alone", and "What is your height, in inches (please type in 510 if your height is 5'10" or 60 if your height is 6 feet) ?" with a text input field.

Height survey

Requester: ##### (anonymized)
HIT Expiration Date: Feb 28 2014, 09:39 PM PST
Reward: \$0.10
Assignments Requested: 200
Description: Height survey
Keywords:

What is your gender?
Male

What is your ethnicity?
White alone

What is your height, in inches (please type in 510 if your height is 5'10" or 60 if your height is 6 feet) ?



How to integrate human and computer computation is a **wide open problem** in programming languages.

Our approach is to use **LINQ** to provide **seamless language extensions**.

But many topics remain unaddressed. How does one program with long-running crowd operations?

How does one combine lazy and eager computation?

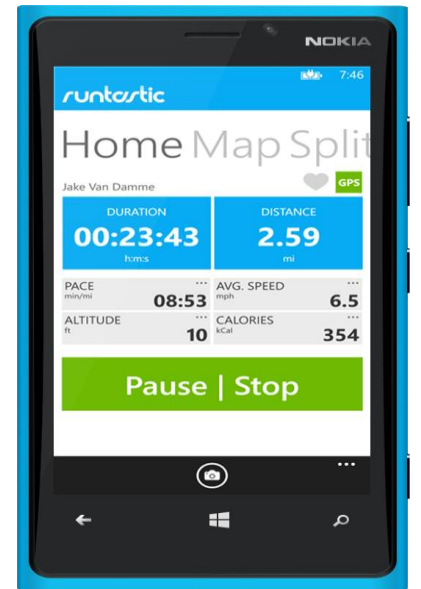




Uncertain $\langle T \rangle$ |

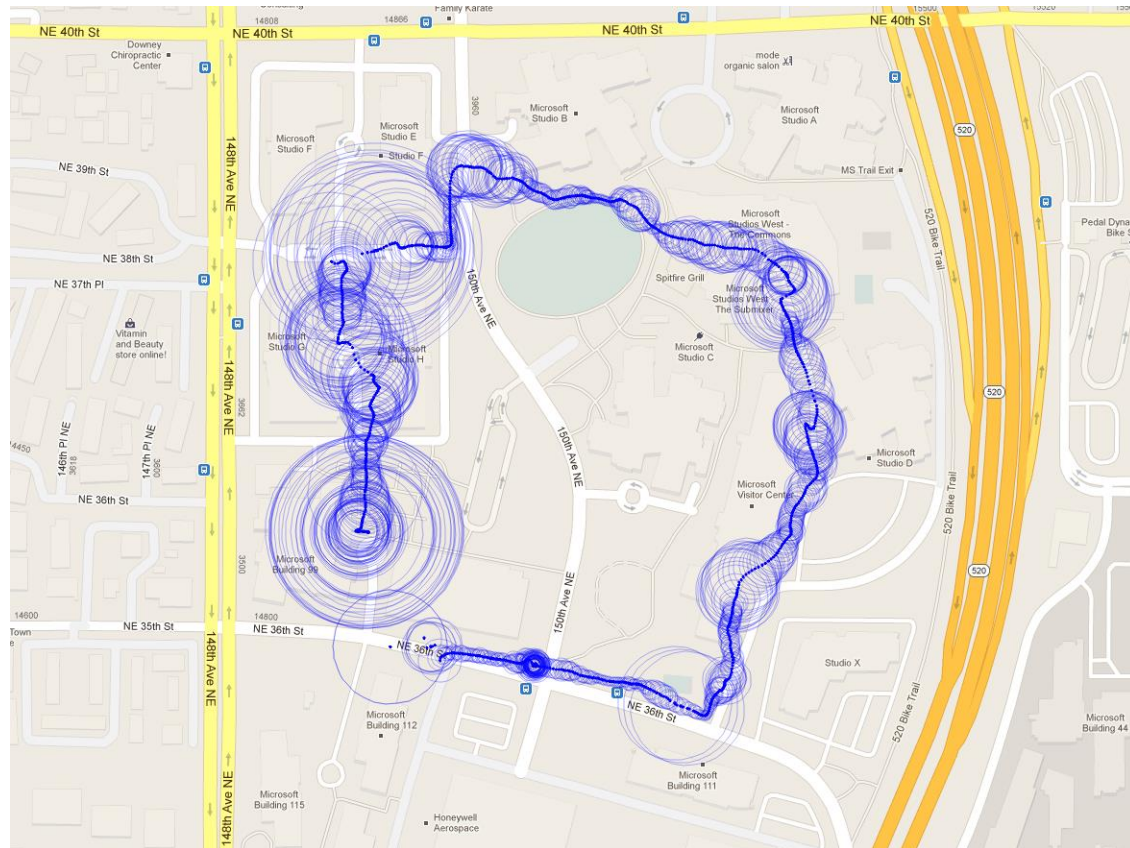


24 mph



59 mph

A WALK WITH A GPS



HOW DO YOU WRITE THIS KIND OF CODE?

```
GeoCoordinate Location;
```

HOW DO YOU WRITE THIS KIND OF CODE?

~~GeoCoordinate Location;~~

Uncertain<GeoCoordinate> Location;



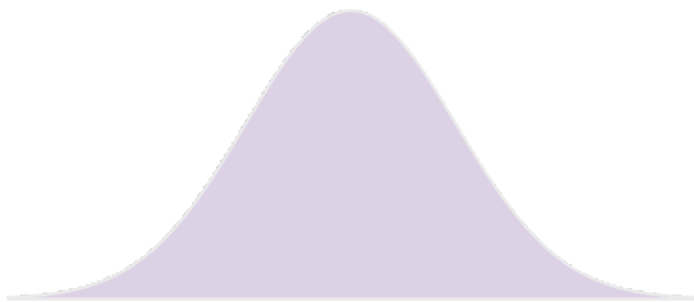
Uncertain<T> is an uncertain type abstraction.

**It encourages developers to
explicitly reason about uncertainty.**

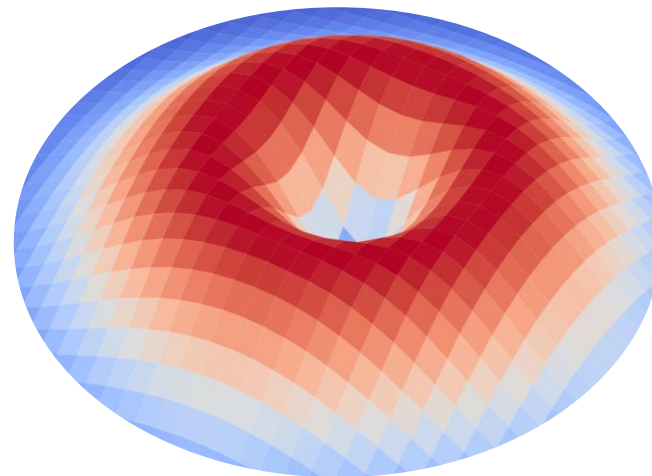
SYNTAX AND SEMANTICS

`Uncertain<GeoCoordinate> LastLoc = GPS.GetLocation();`

A variable of type `Uncertain<T>` is a **random variable**, represented by a **distribution**.



Gaussian distribution



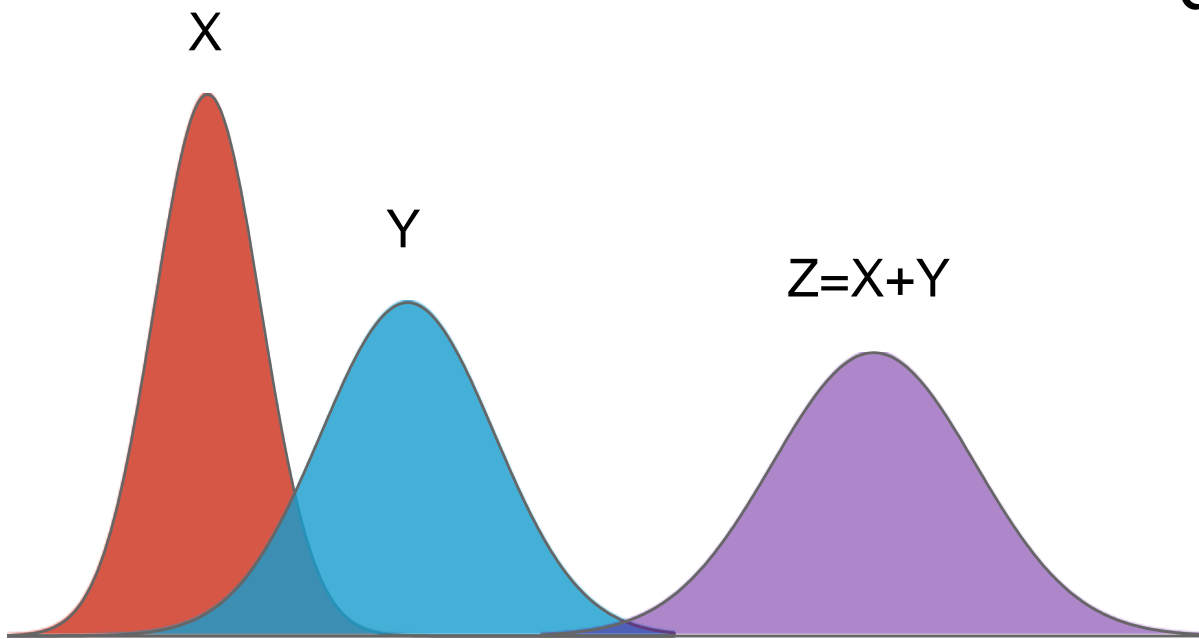
3 D Raleigh Distribution



Arbitrary non-continuous functions

PROGRAMMING WITH `Uncertain<T>`

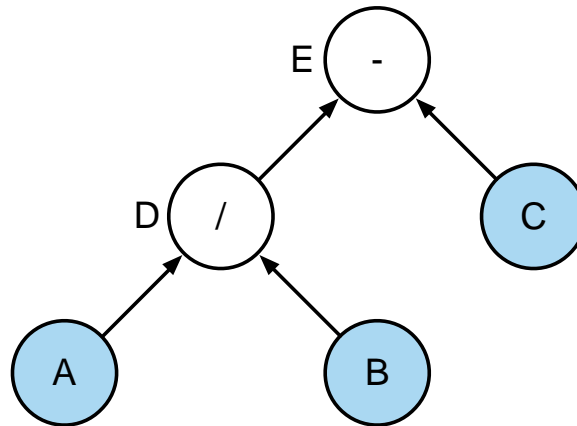
`Uncertain<double> Speed = Dist / 5;` Or more generally, $Z = X + Y$, if X and Y are distributions.



If x is a sample of X
and y is a sample of Y
then $x+y$ is a sample of $X+Y$ *

* if X and Y are independent

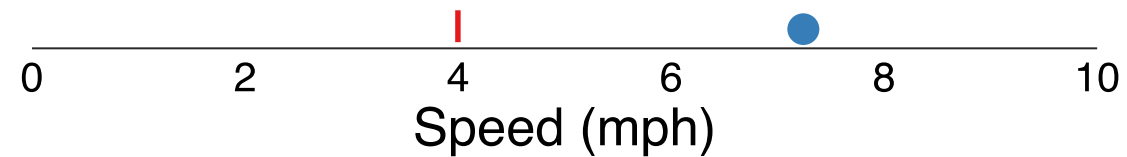
BAYESIAN NETWORK REPRESENTATION:



Sampling function for E recursively samples children.

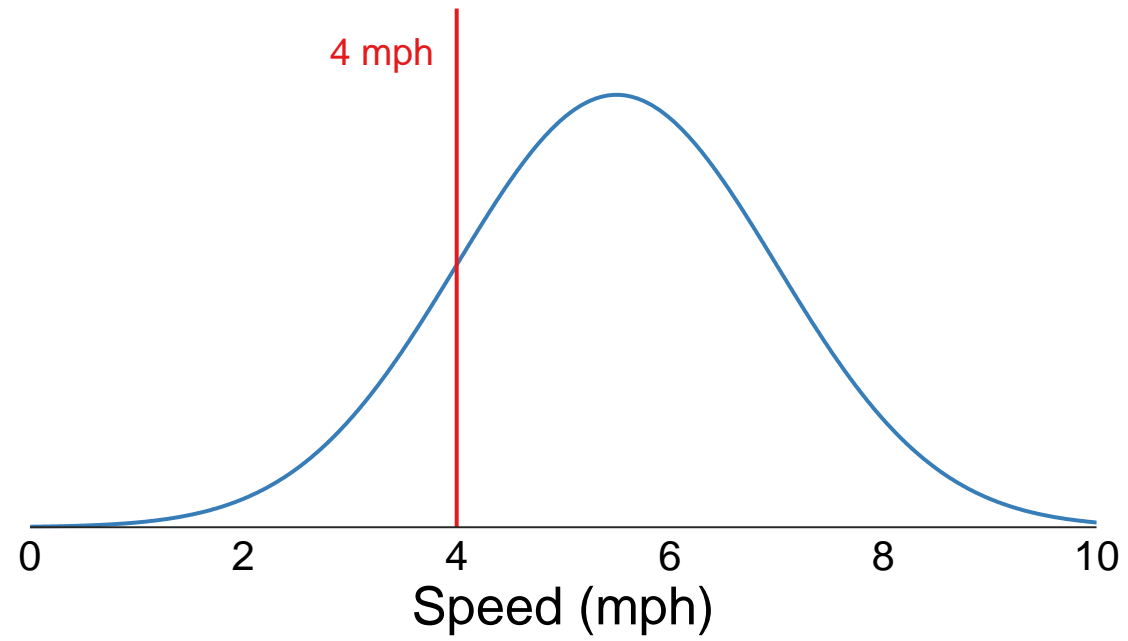
TESTING DISTRIBUTIONS

```
if (Speed > 4) print("Great job!");
```



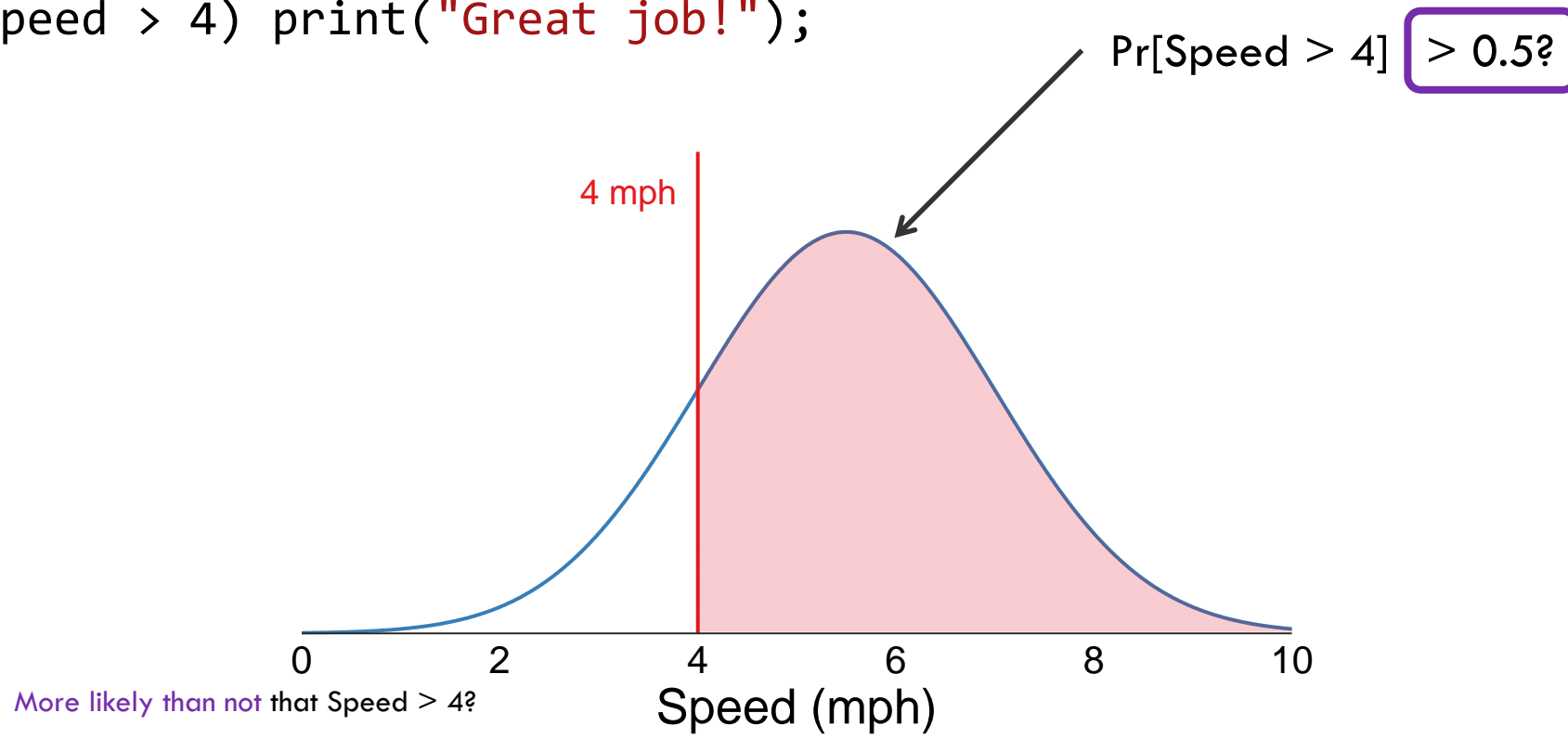
TESTING DISTRIBUTIONS

```
if (Speed > 4) print("Great job!");
```



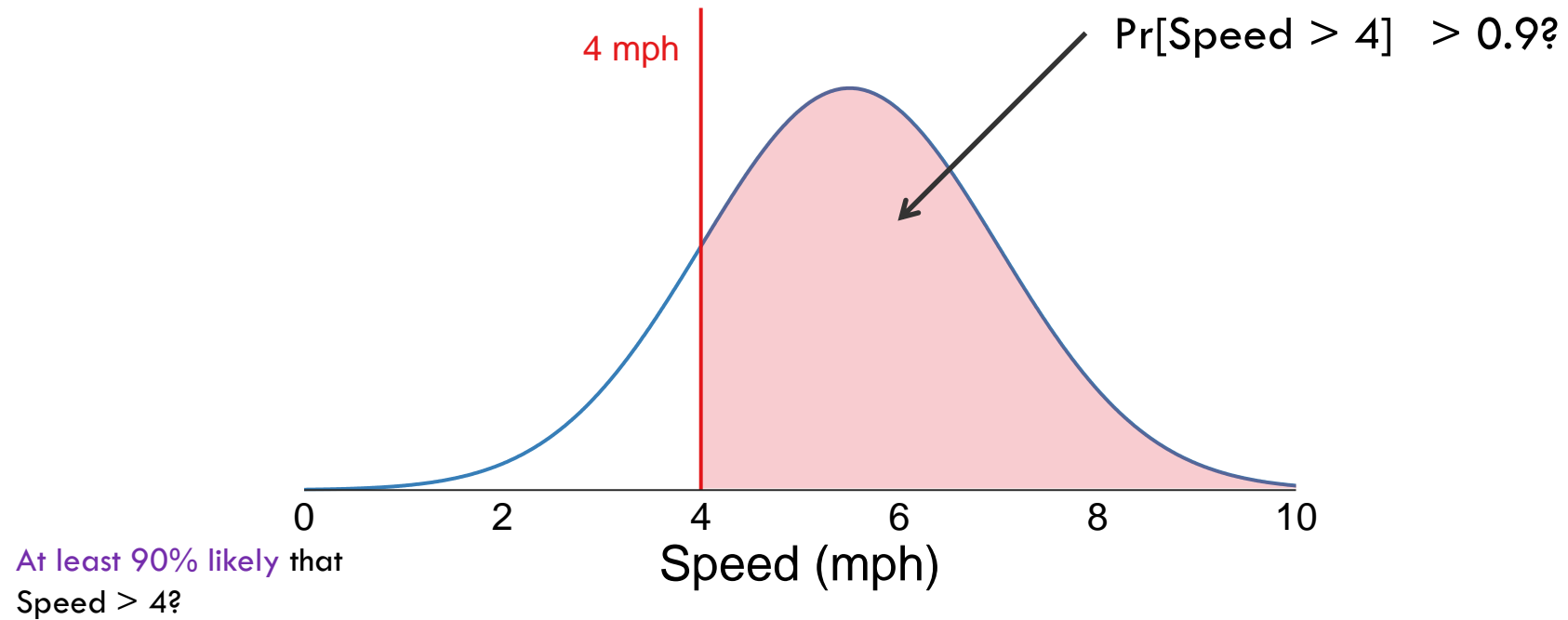
TESTING DISTRIBUTIONS

```
if (Speed > 4) print("Great job!");
```



TESTING DISTRIBUTIONS

```
if (Speed > 4) print("Great job!");
```



TESTING DISTRIBUTIONS

```
if (Speed > 4).Pr(0.9) print("Great job!");
```

null hypothesis $H_0: \Pr[\text{Speed} > 4] \leq 0.9$
 alternate hypothesis $H_A: \underbrace{\Pr[\text{Speed} > 4]}_{\text{approximate!}} > 0.9$

How many samples?

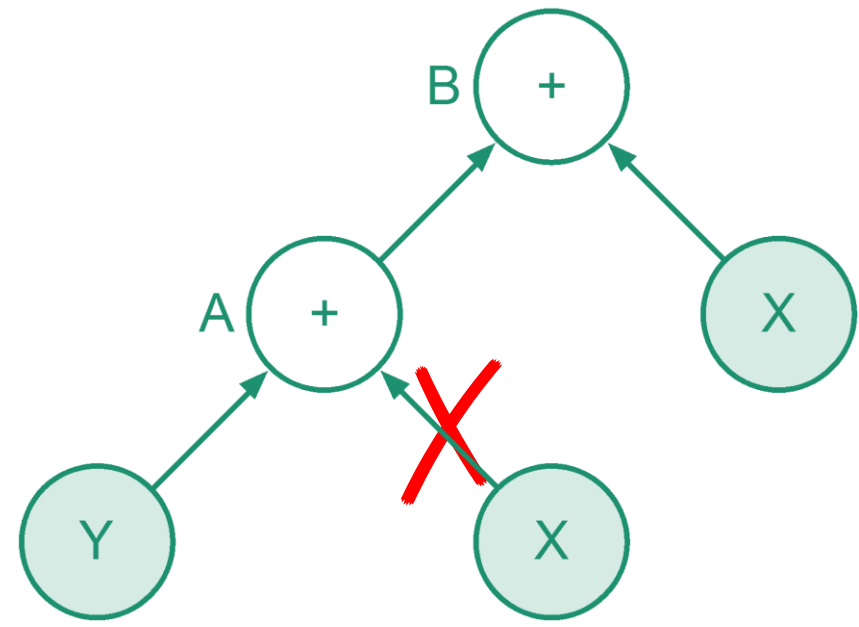
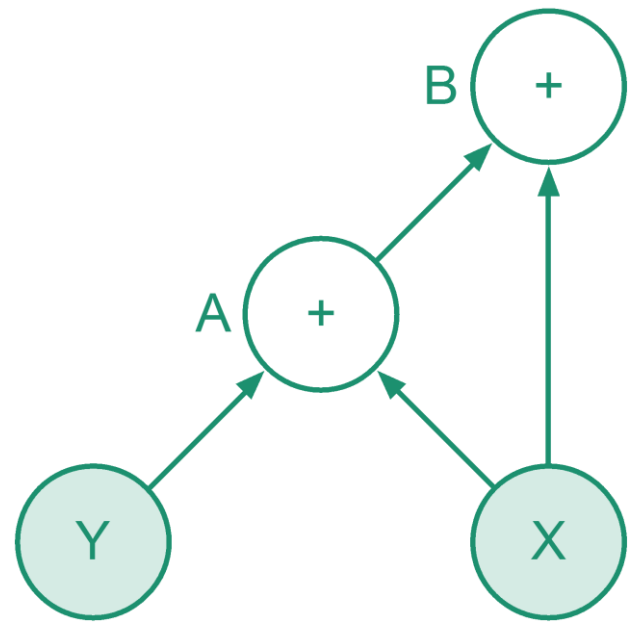
Too many = too slow

Too few = too noisy

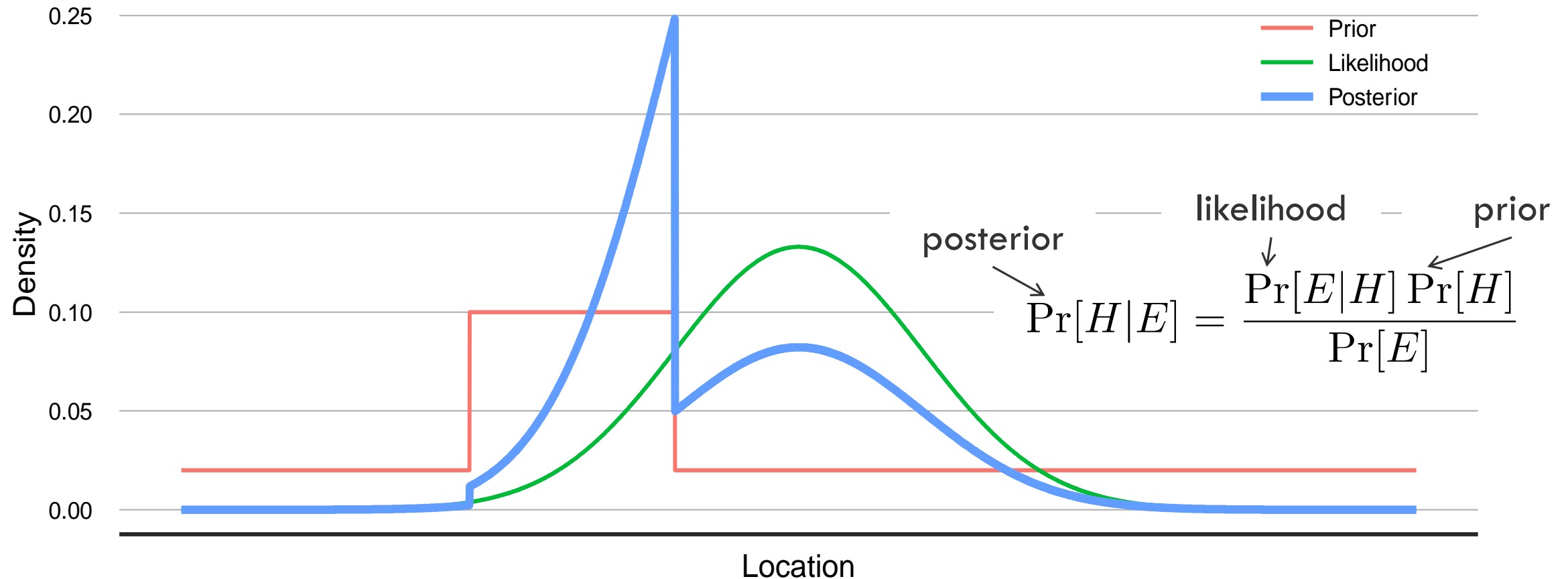
Sequential sampling: sample size depends on progress

BAYESIAN NETWORK ESTIMATION

A and B depend on X – not independent – take one X sample not two!

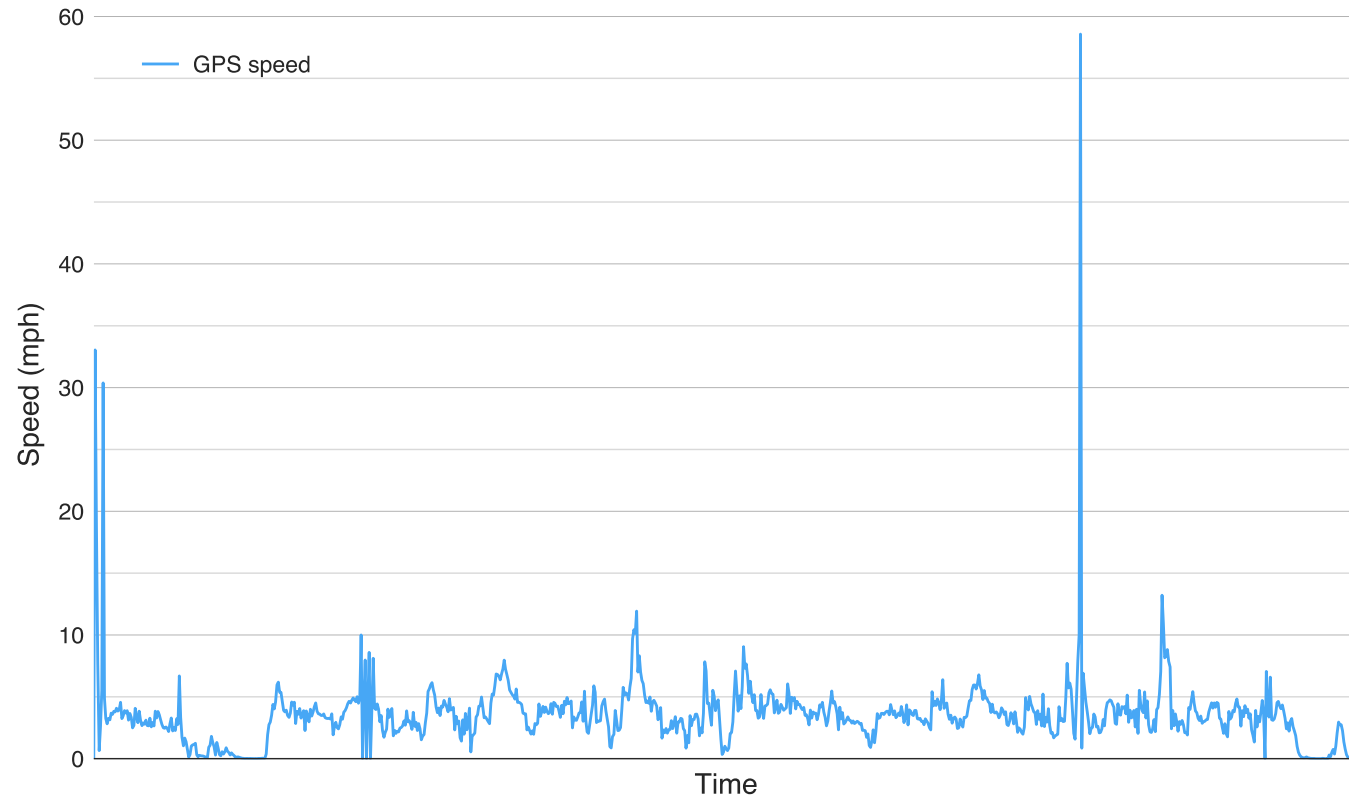


INCORPORATING DOMAIN KNOWLEDGE

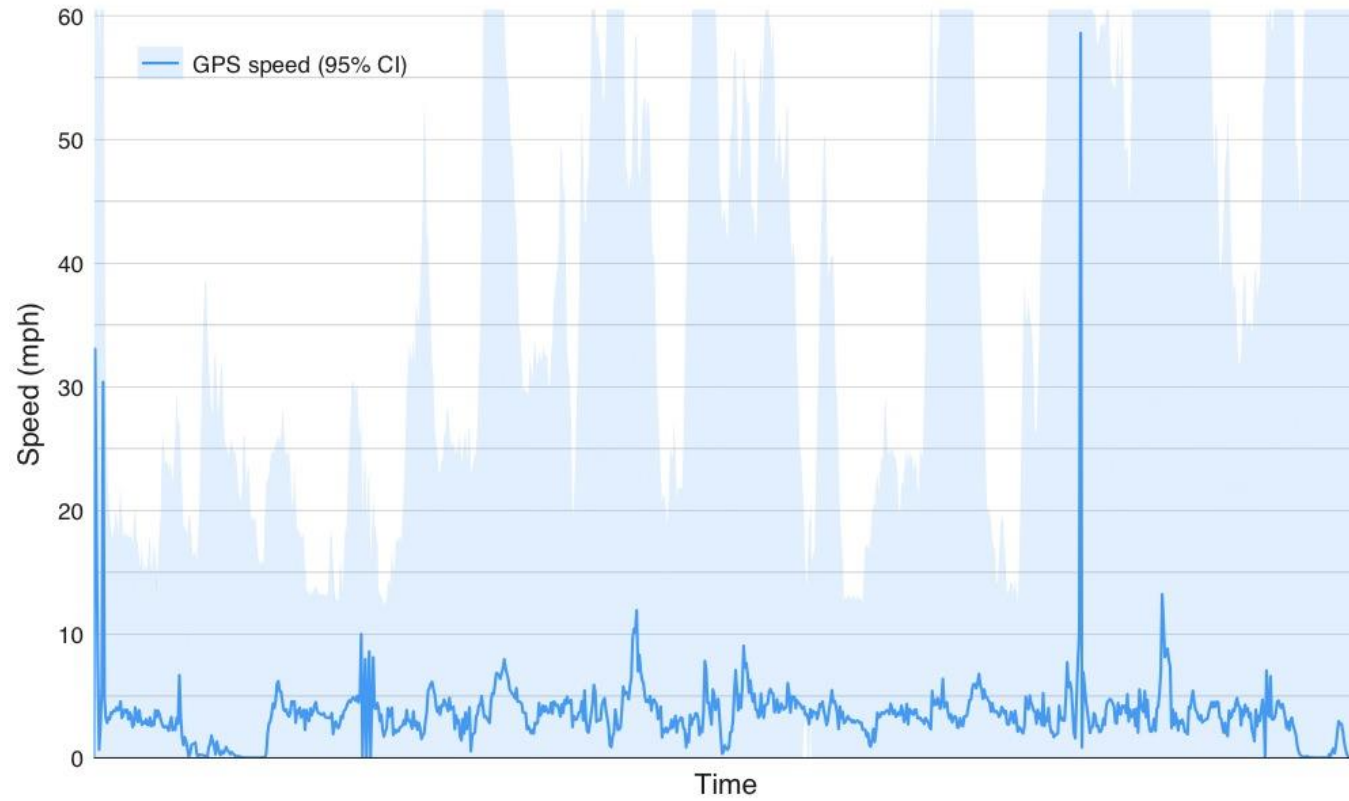


Uncertain<T> language semantics implements Bayes Rule

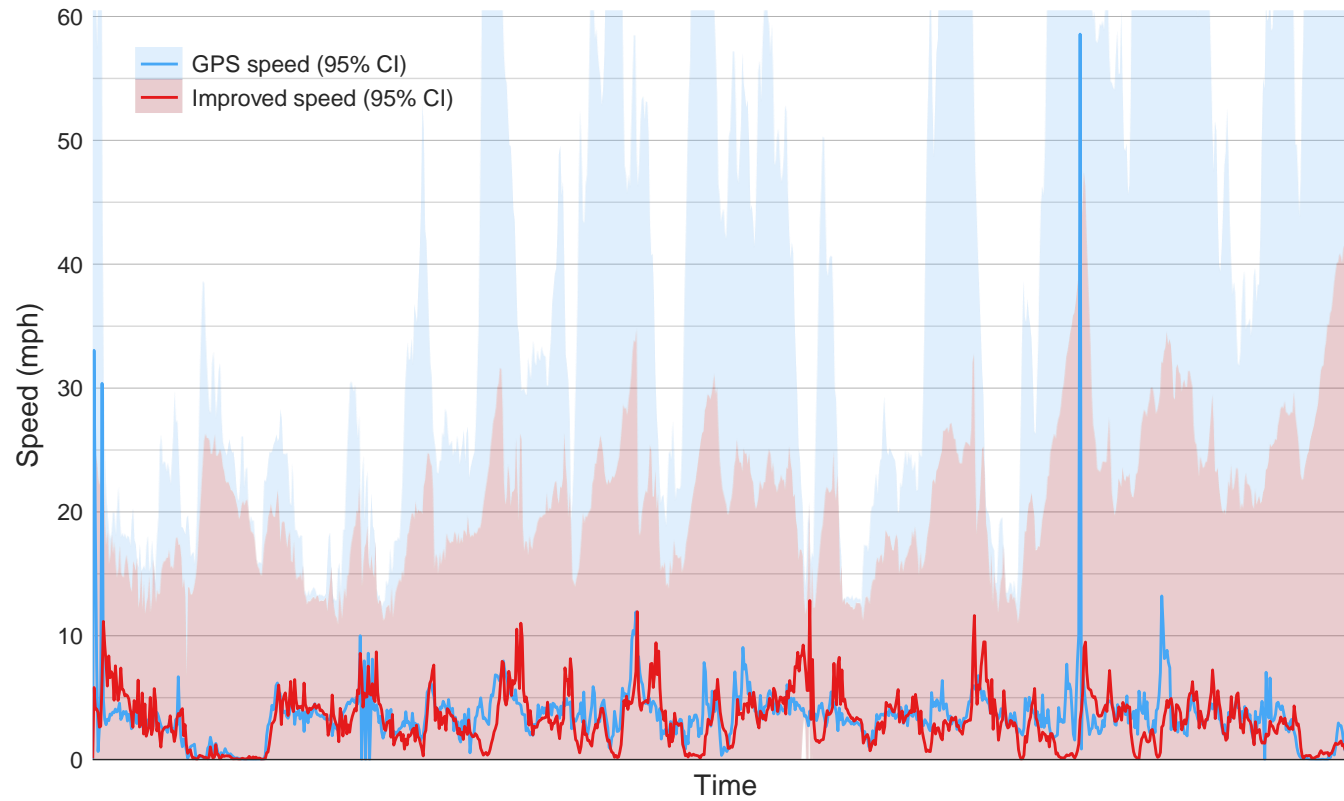
GPS SPEED



GPS SPEED



GPS SPEED



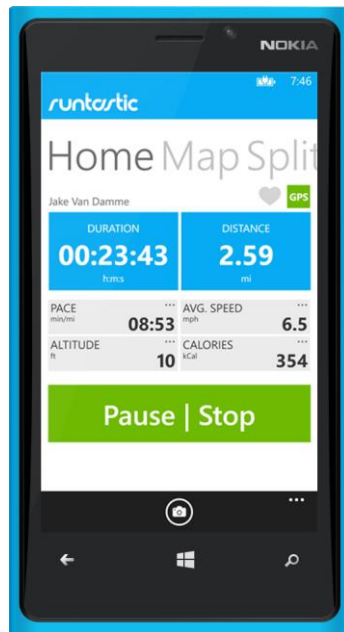


`Uncertain<T>` is an **uncertain type abstraction**.

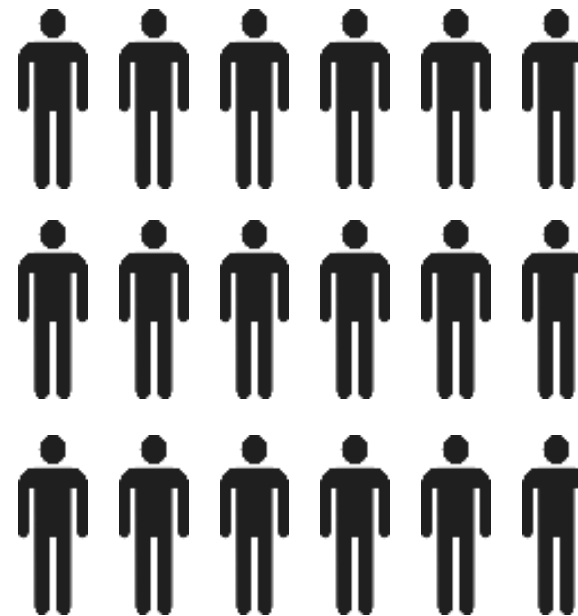
It encourages developers to **explicitly** reason about uncertainty.

BRINGING IT ALL TOGETHER

sensors



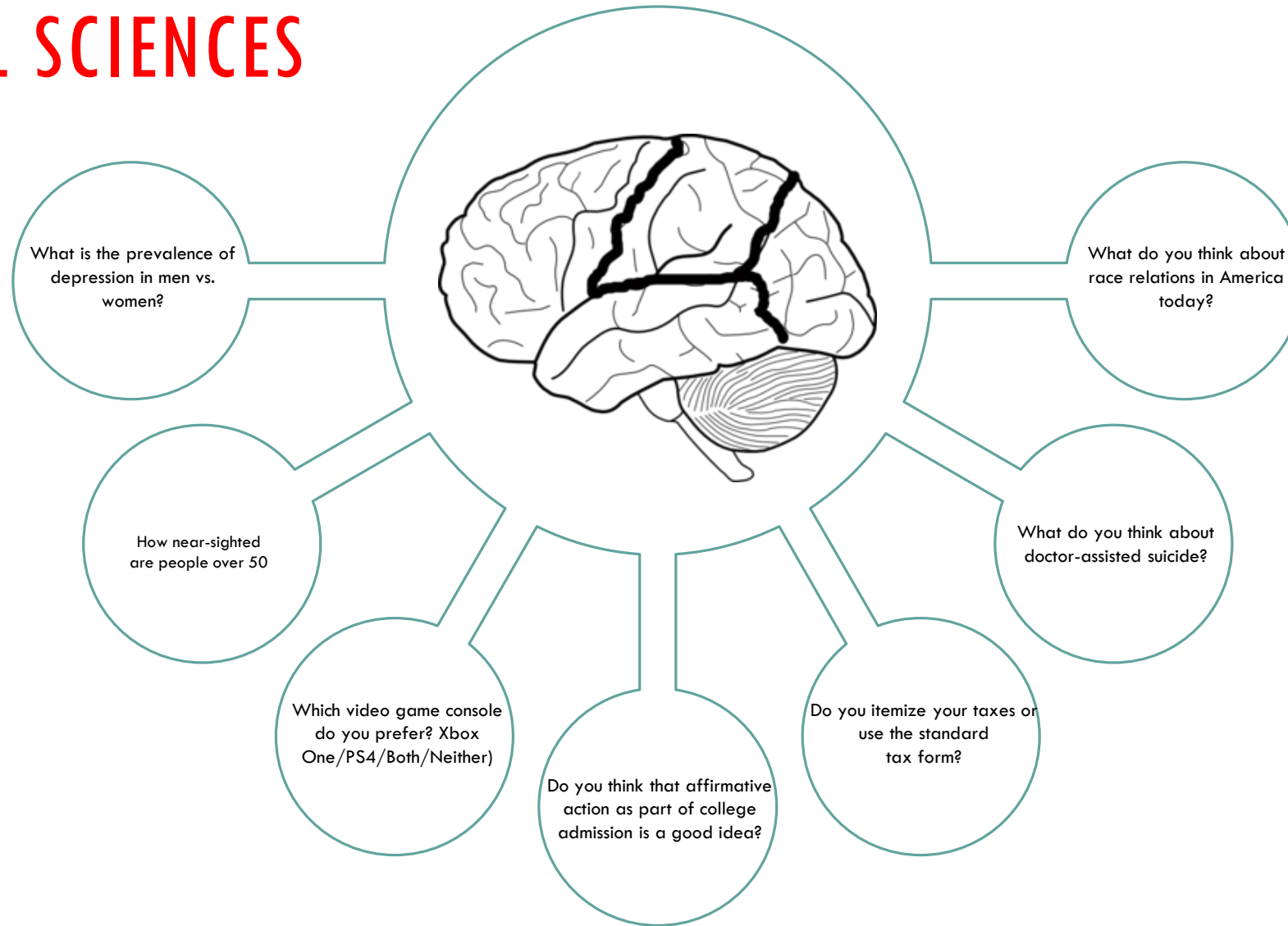
people



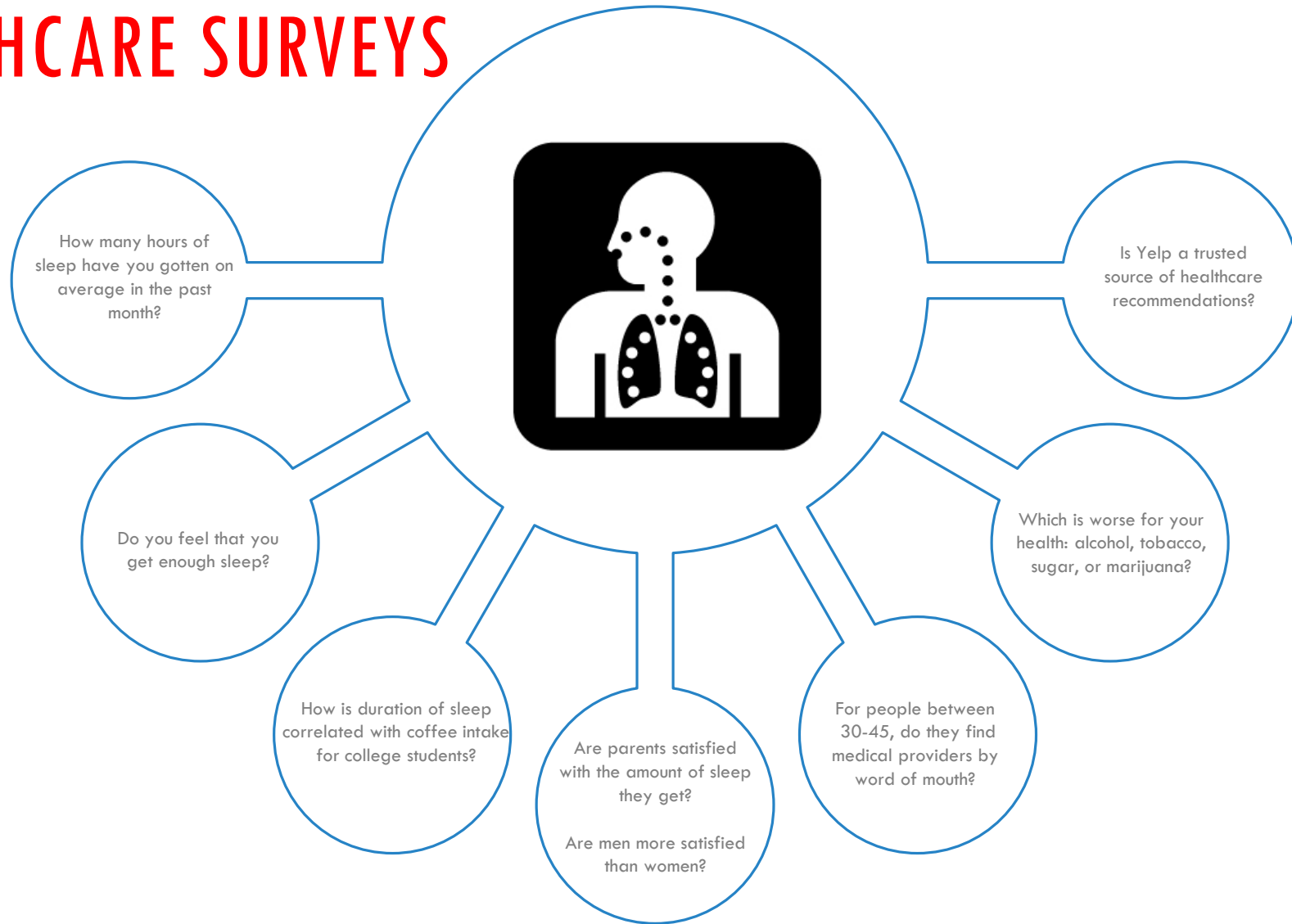


BUILDING INTERPOLL APPLICATIONS

SOCIAL SCIENCES



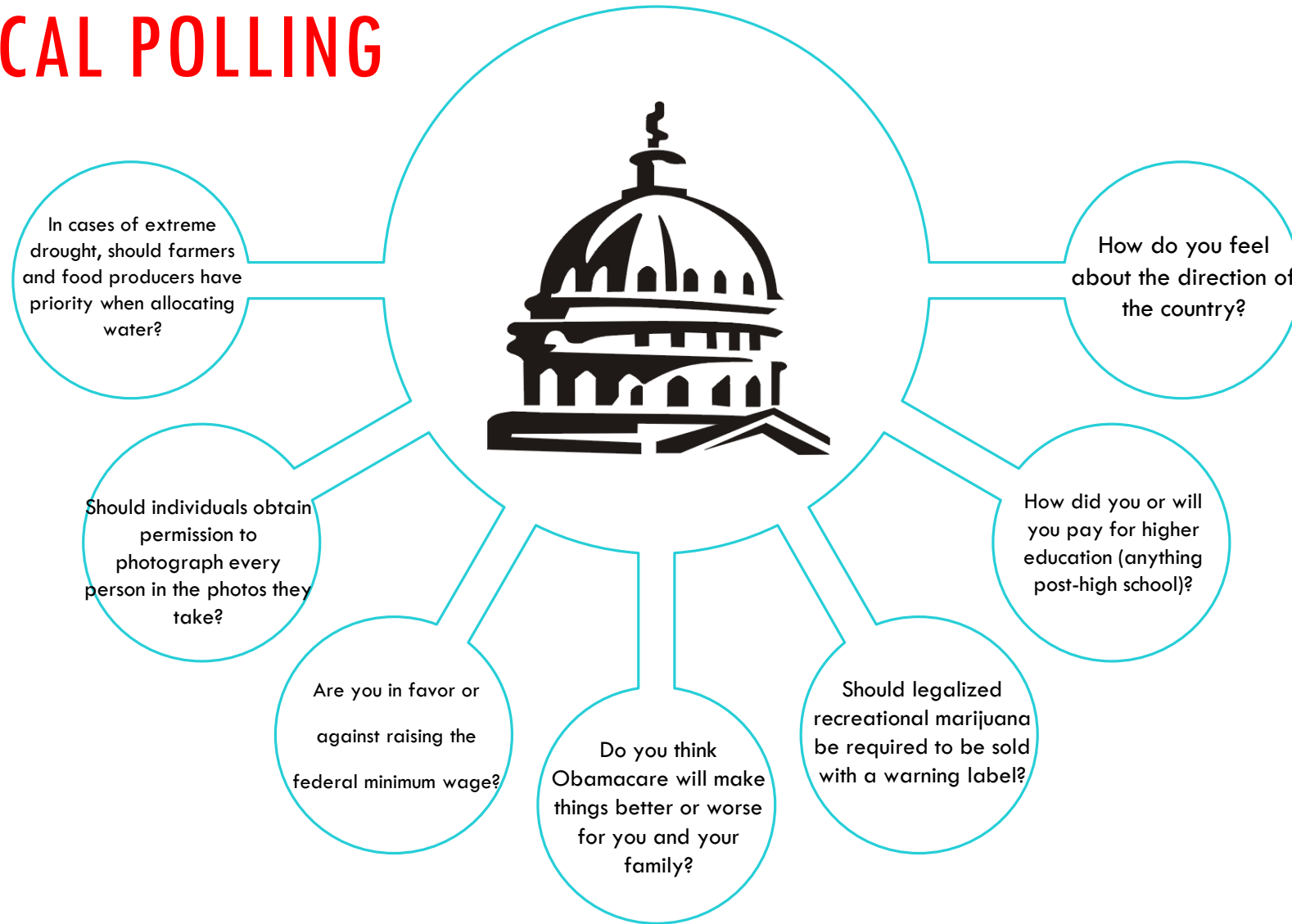
HEALTHCARE SURVEYS



MARKETING



POLITICAL POLLING





Demo: basic surveys

OTHER INTERFACES

Ski Trip Planning

Traffic Center

- Sacramento
- Stockton
- Truckee

Pricing (2012)

Ski Area	Price
Granlibakken Ski Resort	\$24.00
Donner Ski Ranch	\$25.00
Homewood Mountain Resort	\$25.00
Soda Springs	\$36.00
Tahoe Donner	\$41.00
Boreal Mountain Resort	\$49.00
Heavenly Ski Resort	\$61.00
Dodge Ridge Ski Area	\$64.00
Bear Valley Mountain Resort	\$67.00
Sierra at Tahoe	\$77.00
Sugar Bowl Resort	\$77.00
Kirkwood	\$79.00
Northstar	\$91.00
Alpine Meadows Ski Area	\$96.00
Squaw Valley USA	\$99.00

Most Recent Road Conditions

Traffic Center	Location	Time	Traffic Inc Type	Location Desc.
Sacramento	Jefferson Blvd / Babel Slough Rd	03:15 AM	Car Fire	NB JNO 3/4 MILE
Sacramento	Main St	05:18 AM	Assist with Construction	RIO VISTA PD
Sacramento	Mm50 / Still Meadow Rd	03:02 AM	Traffic Hazard	WB
Sacramento	Watt Ave / Don Julio Blvd	08:58 PM	SILVER Alert	FREWILL BAPTIST CHURCH
Sacramento	Watt Ave / S Watt Ave	03:19 PM	SILVER Alert	WATT AVE IN SACRAMENTO
Stockton	3330 N Ad Art Rd	03:19 PM	SILVER Alert	STOCKTON CHP OFFICE
Truckee	10077 Frates Ln	03:15 PM	SILVER Alert	TRUCKEE CC
Truckee	180 E / 180 E Rainbow Rd Ofr	04:25 PM	CLOSURE of a Road	

Power View Fields

- ACTIVE | ALL
- FactPricing
- Ski Areas
- TrafficCenters

Edit Survey

Jam Session Planning

Let's get together to make some music. Let me know when you're free and what you play!

What's your name?

What instrument do you play?

When's best for you?

+ Add New Question

EDIT QUESTION

Question: What's your name?

Question Subtitle:

Response Type: Text

Required:

Default Answer:

Done Delete Question

Share Survey Save and View Close

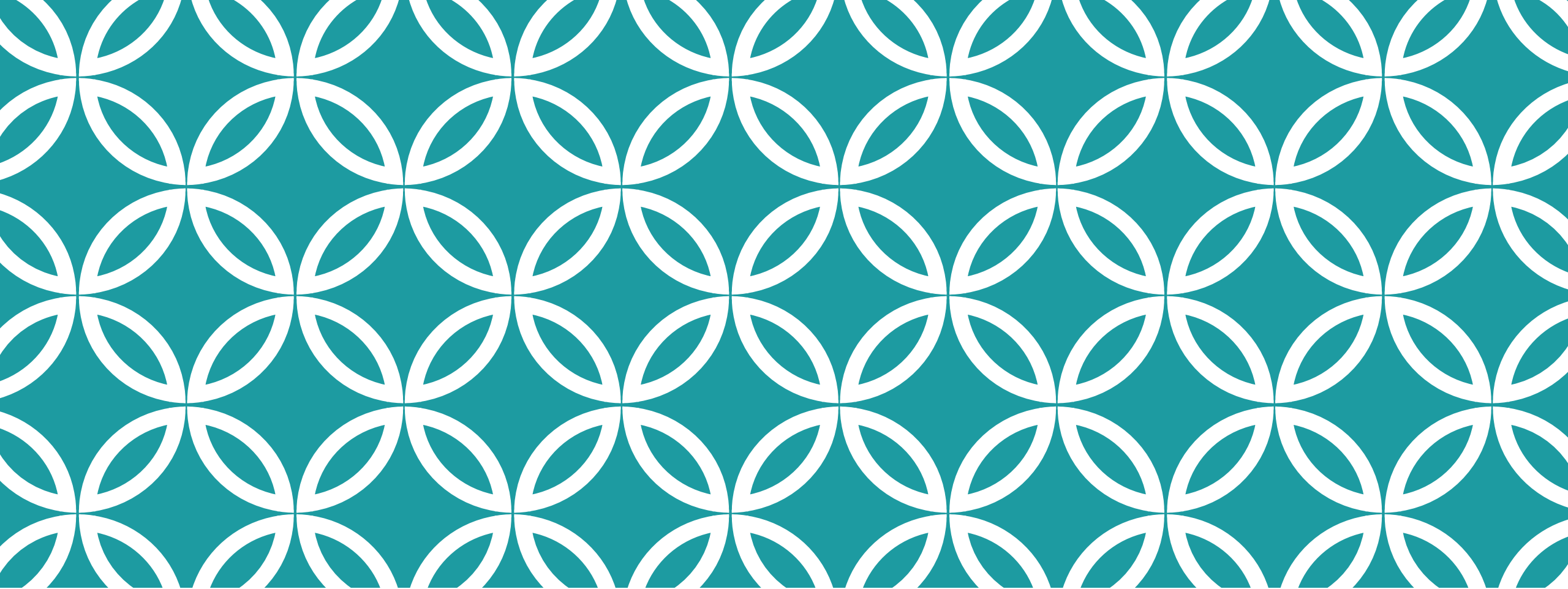


FUNDAMENTAL PROBLEMS



MOVING AWAY FROM SMALL AND UNREPRESENTATIVE SAMPLES





POWER ANALYSIS

POWER ANALYSIS

Determine the number of samples for a query

We can sample from the crowd sequentially until we satisfy or disprove our hypothesis.

We will poll the crowd for more until our stopping criterion is reached.

The stopping criterion allows us to conclude that the hypothesis can be proven or disproven with the required level of confidence.

EXAMPLE QUESTION: HEIGHT

```

var people = GetPeople(GetDescription(), 200, false, false);
var height = from person in people
              select new
              {
                  Height = person.PoseQuestion<int>(
                      "What is your height, in centimeters " +
                      "(if you know your height in inches, you can convert to
                      "centimeters using a calculator here: http://www.calculator.net
                      "Please be careful when typing in your height. Invalid input will result in an error message.
                      ),
                  Gender = person.Gender,
                  Ethnicity = person.Ethnicity,
              };

var males = from person in height
            where person.Gender == Gender.MALE
            select person.Height;
var females = from person in height
              where person.Gender == Gender.FEMALE
              select person.Height;

if (males.ToRandomVariable(false) > females.ToRandomVariable(false))
{
    Console.WriteLine("Males are taller than females.");
}

```

N=29

Once we remove the outliers

height = from person in height where

where

person.Height >= 140 &&

person.Height <= 220)

N=27

ARE THE TWO CORRELATED?

```
var happinessData = from person in people
                    select new
                    {
                        Consider = person.PoseCodedQuestion(
                            "I feel tense or 'wound up'", new Tuple<string,
                        Enjoy = person.PoseCodedQuestion(
                            "I still enjoy the things I used to enjoy", new
                        Awful = person.PoseCodedQuestion(
                            "I get a sort of frightened feeling as if somet
                        Laugh = person.PoseCodedQuestion(
                            "I can laugh and see the funny side of things",
                        Worry = person.PoseCodedQuestion(
                            "Worrying thoughts go through my mind", new Tup
                        Cheerful = person.PoseCodedQuestion(
                            "I feel cheerful", new Tuple<string, int>("Not
                        Sit = person.PoseCodedQuestion(
                            "I can sit at ease and feel relaxed", new Tuple
                        Slowed = person.PoseCodedQuestion(
                            "I feel as if I am slowed down", new Tuple<stri
                        Frightened = person.PoseCodedQuestion(
                            "I get a sort of frightened feeling like 'butte
```

```
var scores = from data in happinessData
              select new
              {
                  Anxiety =
                      data.Consider + data.Awful + data.Worry +
                      data.Sit + data.Frightened + data.Restless + data.Panic,
                  Depression = data.Enjoy + data.Laugh + data.Cheerful + data.Slowed +
                      data.LostInterest + data.LookForward + data.GoodBook,
                  Gender = data.Gender,
                  Income = data.Income,
                  Education = data.Education,
                  Ethnicity = data.Ethnicity,
                  Employment = data.Employment,
              };
```

Hospital anxiety and depression scale (HADS)

Normal 0-7

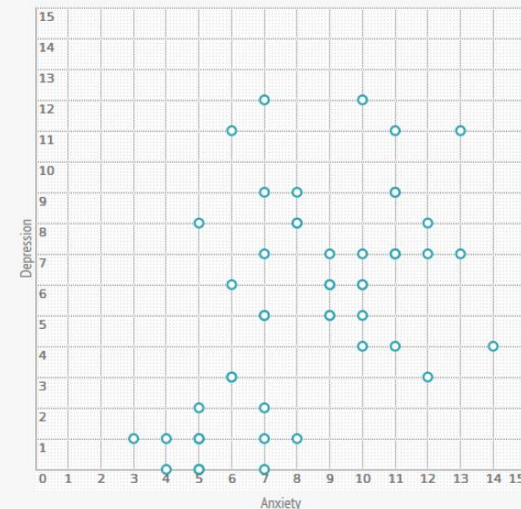
Mild 8-10

Moderate (11-14)

Severe (15-21)

RAW DATA FROM MTURK

ANXIETY VS. DEPRESSION



DOES MONEY BUY HAPPINESS (OR AT LEAST TRANQUILITY)?

```
var rich = from person in scores
           where
               person.Income == Income.INCOME_35_000_TO_49_999 ||
               person.Income == Income.INCOME_75_000_AND_OVER ||
               person.Income == Income.INCOME_50_000_TO_74_999
           select person.Anxiety;

var poor = from person in scores
           where
               person.Income == Income.INCOME_1_TO_4_900 ||
               person.Income == Income.INCOME_10_000_TO_14_999 ||
               person.Income == Income.INCOME_15_000_TO_24_999
           select person.Anxiety;

if (rich.ToRandomVariable(false) < poor.ToRandomVariable(false))//(p >
{
    Console.WriteLine("Rich are more anxious than poor: " + "\tYes");
}
else
{
    Console.WriteLine("Rich are more anxious than poor: " + "\tNo");
}
```

Are rich **more anxious** than poor?

N = 105

expected value for poor=8.5714,
expected value for rich=7.9619

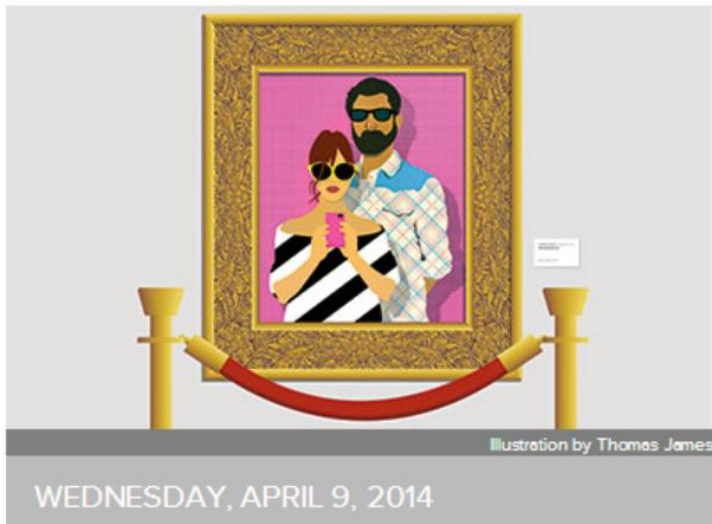
DEBATES: INTELLIGENCE SQUARED

MILLENNIALS DON'T STAND A CHANCE

DEBATE DETAILS

THE PANEL

RESULTS



Task	Outcome	Power	Cost
MilennialsDontStandAChance	No	37	\$3.70
MinimumWage	No	43	\$4.30
RichAreTaxedEnough	No	51	\$5.10
EndOfLife	No	53	\$5.30
BreakUpTheBigBanks	Yes	73	\$7.30
StrongDollar	No	85	\$8.50
MarginalPower	No	89	\$8.90
GeneticallyEngineeredBabies	Yes	135	\$13.50
AffirmativeActionOnCampus	Yes	243	\$24.30
ObesityIsGovernmentBusiness	No	265	\$26.50

CONVERGENCE CURVES: SEQUENTIAL PROBABILITY RATIO TEST (SPRT) OR WALD, 1945

Sequential probability ratio test: To implement this, we build a sequential acceptance plan. Let $H_0 : p + \epsilon$ and $H_A : p - \epsilon$ where $p = 0.5$ by default and can be overloaded by a programmer. `Uncertain(T)` calculates the cumulative log-likelihood ratio for each sample:

$$\Delta_L = k \log(H_A/H_0) + (n - k) \log(H_0/H_A)$$

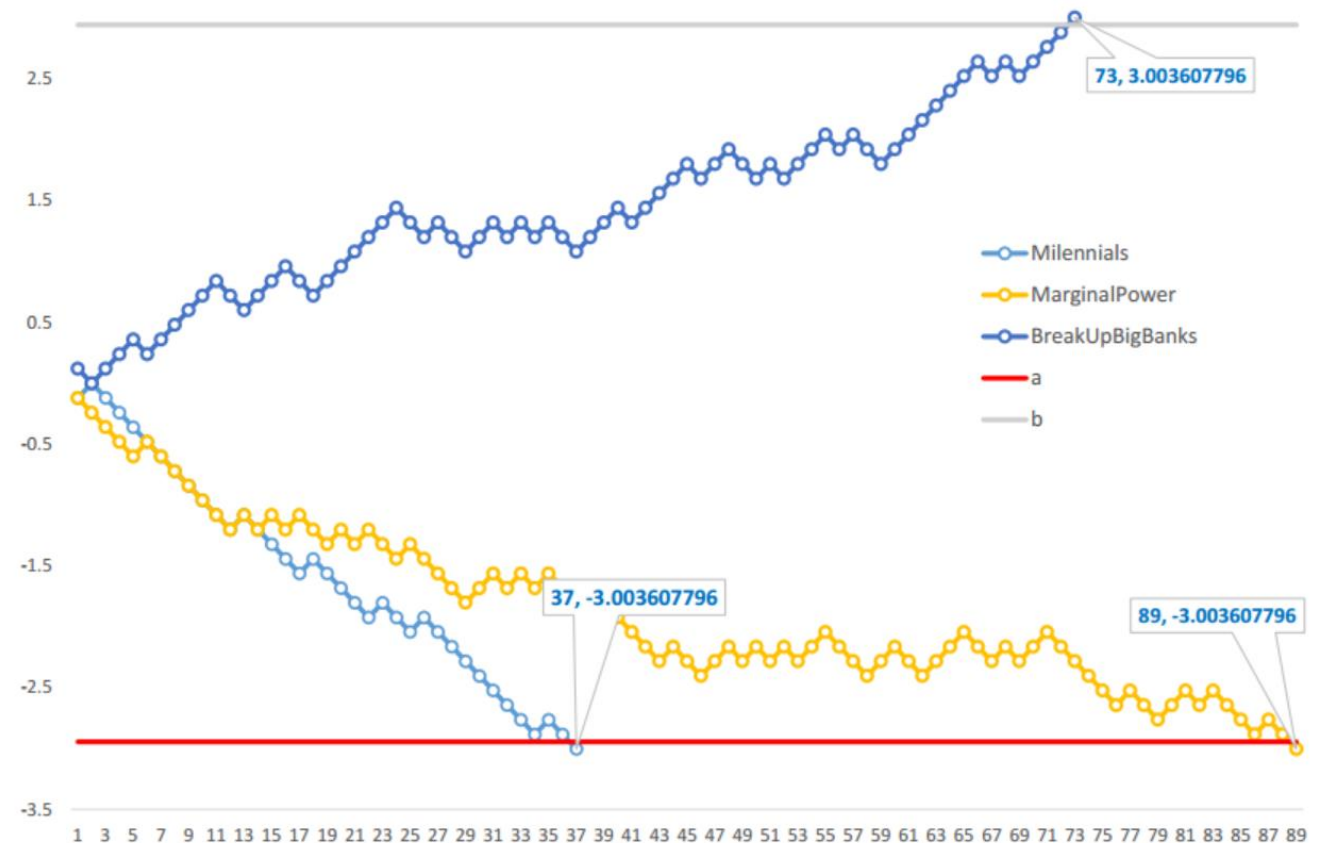
where n is the number of samples taken thus far and k is the number of successes out of those n trials. If

$$\Delta_L \leq \log(\alpha/(1 - \alpha)) \quad =a$$

then `Uncertain(T)` evaluates the conditional as false while if

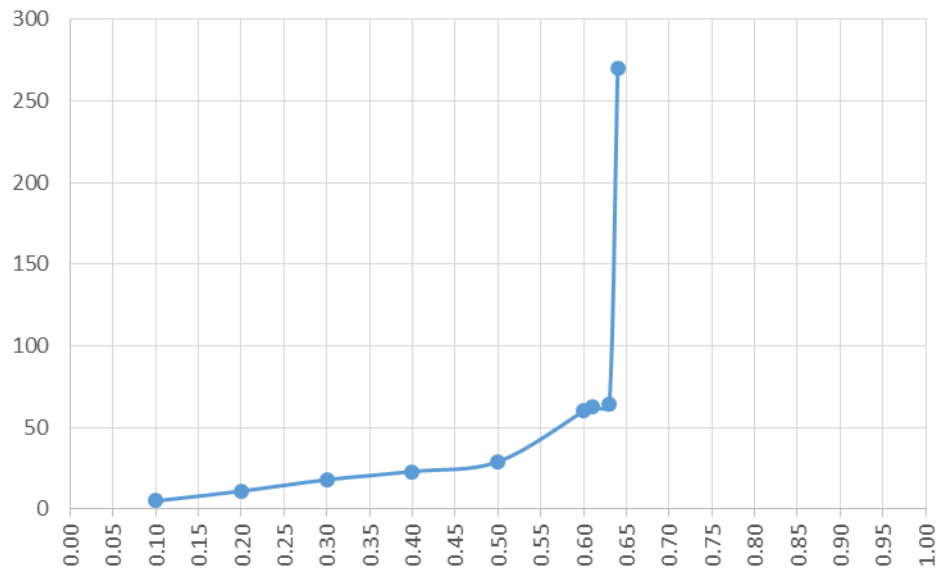
$$\Delta_L \geq \log((1 - \alpha)/\alpha) \quad =b$$

the conditional is true.

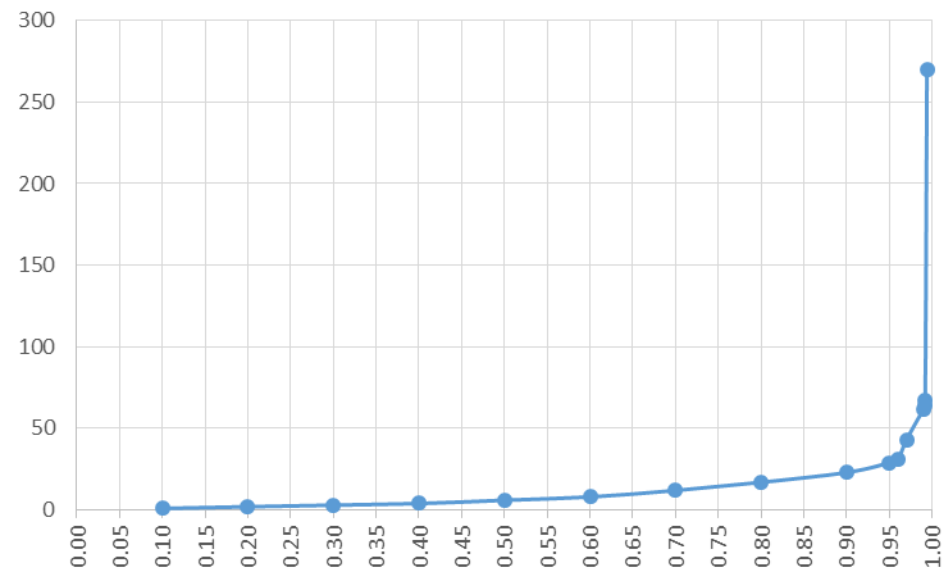


BE CAREFUL WITH YOUR PARAMETERS

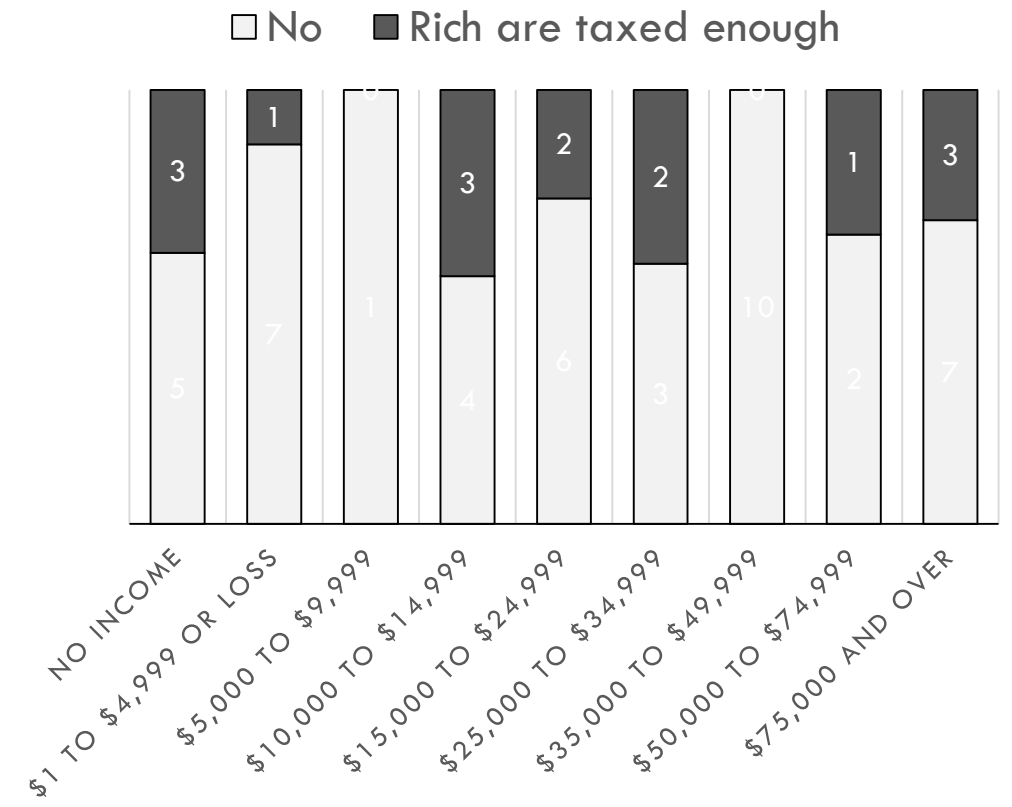
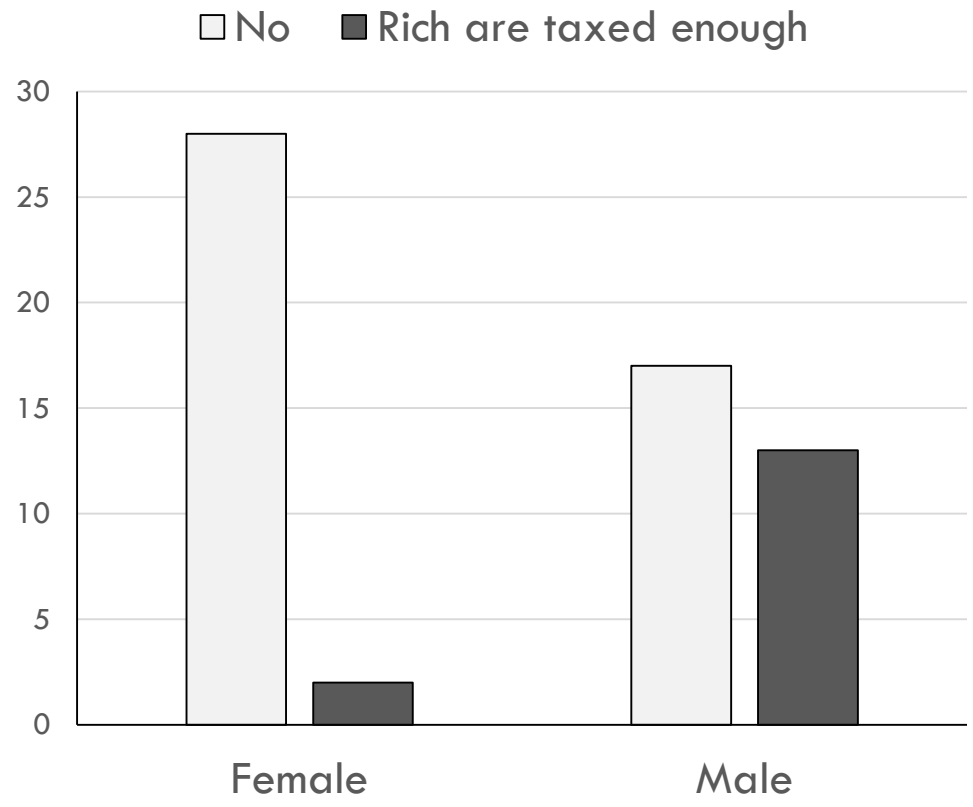
As we vary the probability



As we vary the confidence

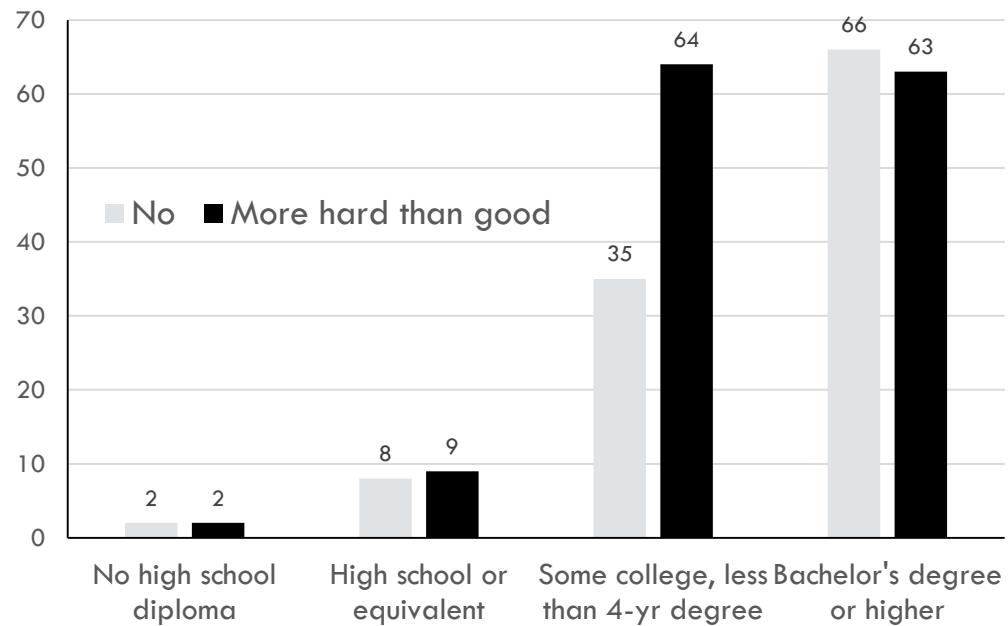


TAXATION, BY GENDER AND INCOME

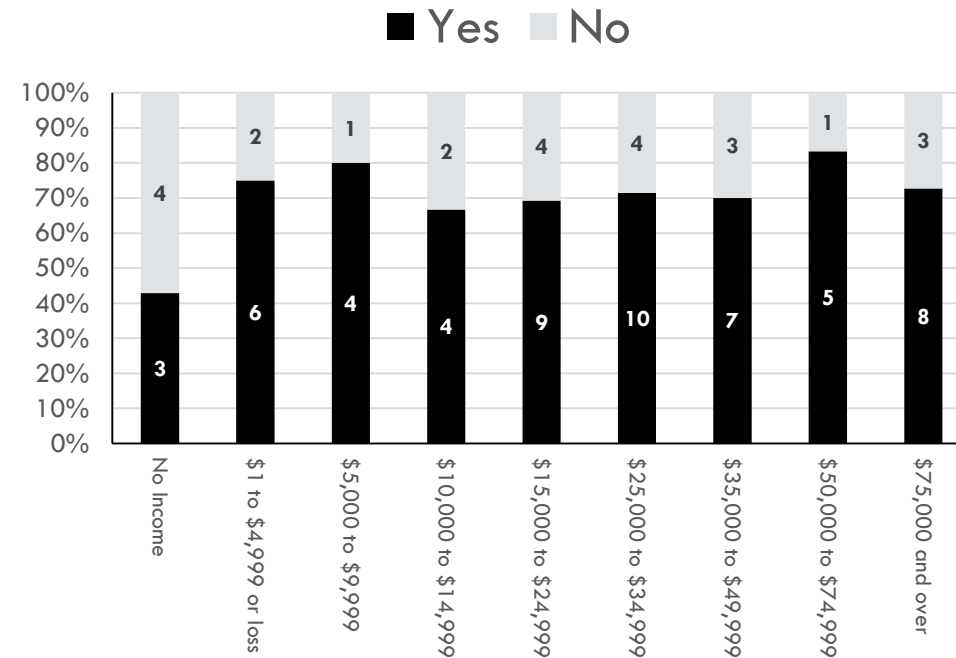


TWO MORE CONUNDRUMS

Affirmative action does more harm than good

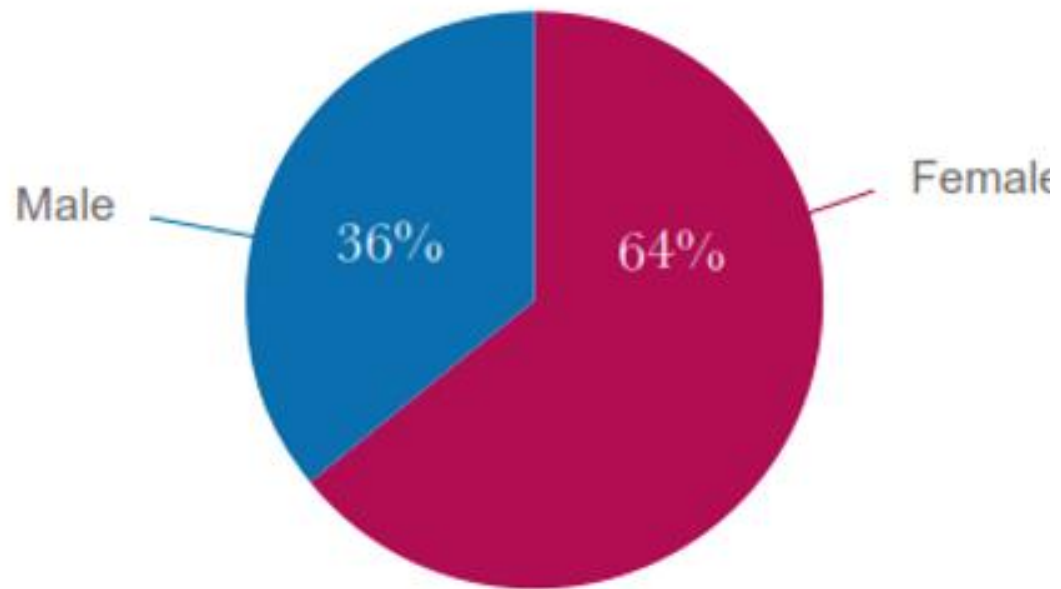


We should break up the big banks

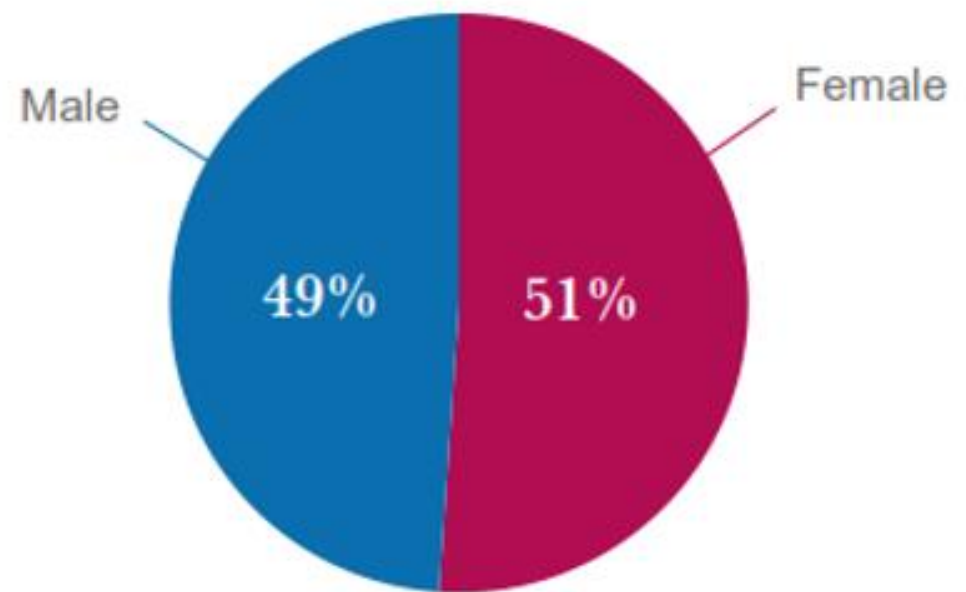


PRIORS FOR THE CROWD

Instant.ly crowd

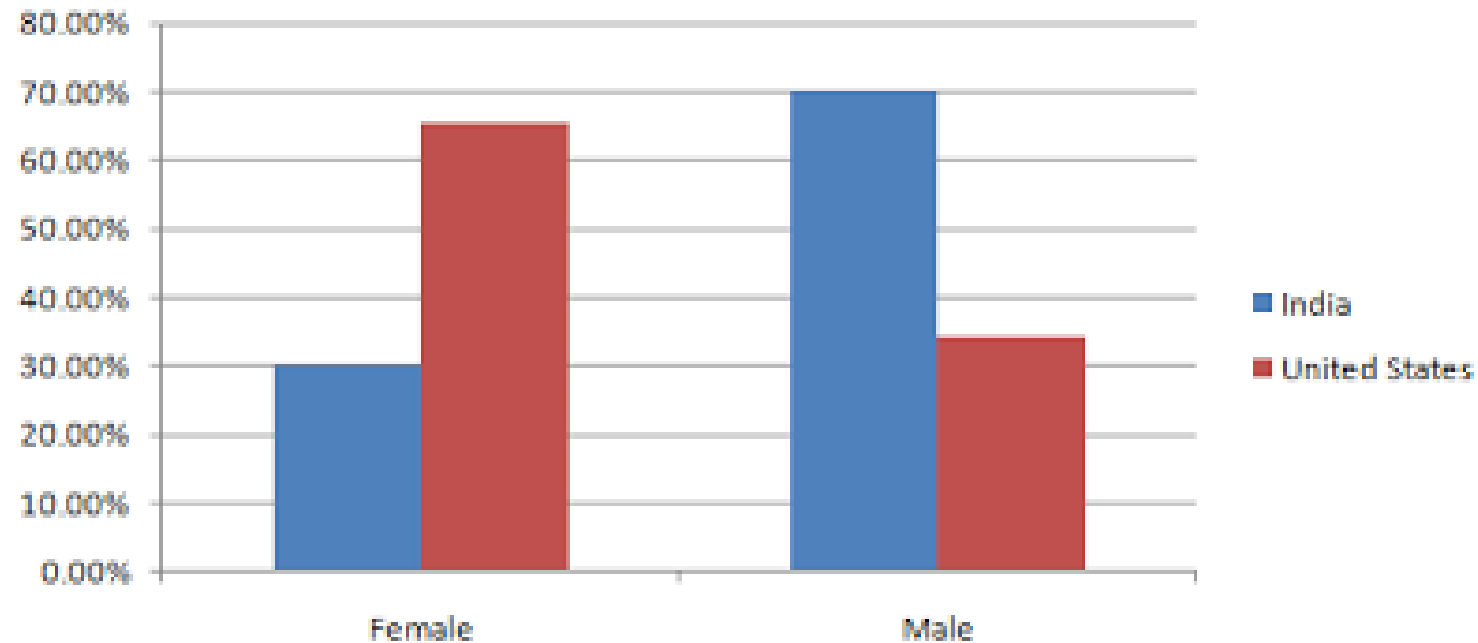


US census



SEGMENTS OF THE CROWD LOOK DIFFERENT

Gender Breakdown



THE UNBIAS OPERATOR

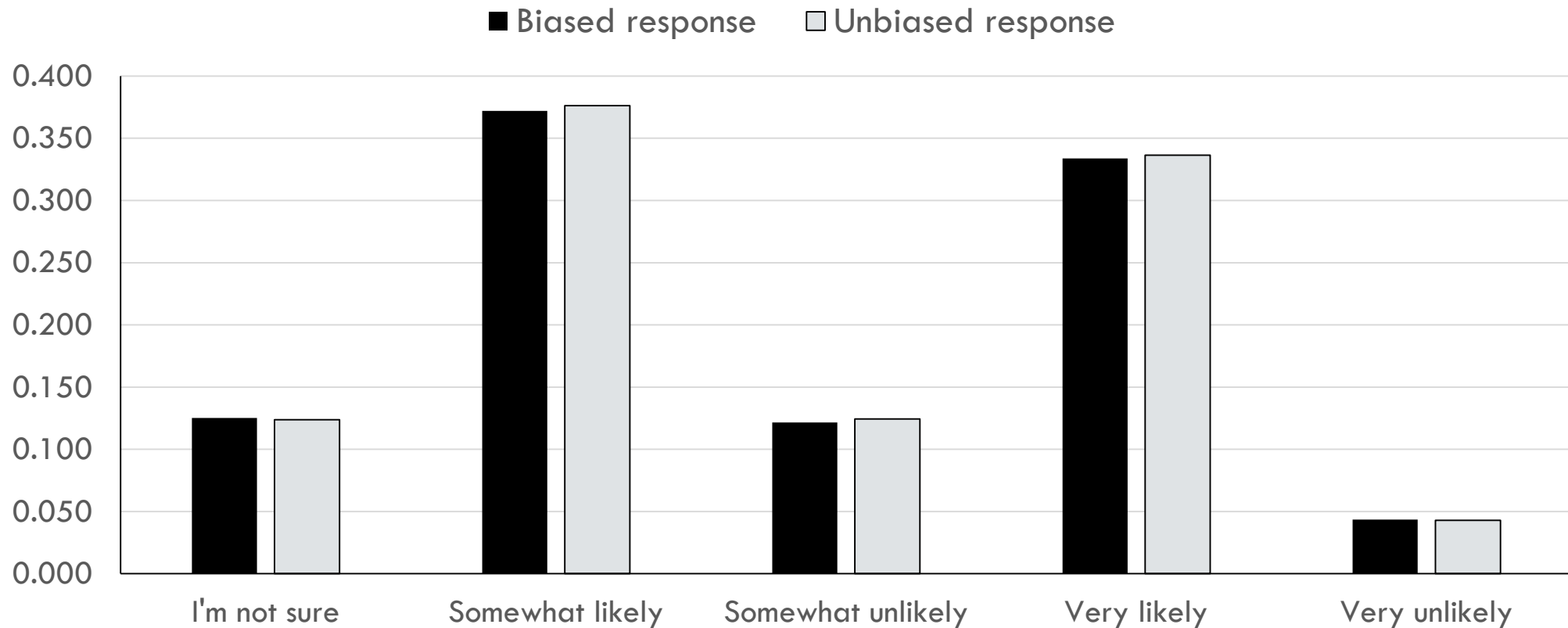
```

var photoAttitudes = (from person in people
    select new
    {
        Used = person.PoseQuestion<bool>(
            "Have you ever hired a professional photograph
        WorthIt = person.PoseQuestion<bool>(
            "Do you feel the money you spent was worth the
        Quality = person.PoseQuestion(
            "How would you rate the quality of the pictures
            "\u2605", "\u2605\u2605", "\u2605\u2605\u2605",
        HowLikely = person.PoseQuestion(
            "How likely are you to
            "Very likely", "Somewh // priors for demographics.
        var mturk = MTurkPriors.DefaultPriors;
        var census = CSPSlicedPriors.DefaultPriors;

        var correctedAttitudes = Unbiasing.Unbias(photoAttitudes,
            p => p.HowLikely,
            p => p.Gender, mturk, census);

```

UNBIASING IN INTERPOLL



LOCAL OPTIMIZATIONS

```
var males = from person in height
            where person.Gender == Gender.MALE
            select person.Height;

var females = from person in height
              where person.Gender == Gender.FEMALE
              select person.Height;
```

```
var unemployed =
    from person in people
    where person.Employment == Employment.WORKED_FULL_TIME_YEAR_F
    select person;

foreach (var person in unemployed)
{
    if (person.Employment == Employment.DID_NOT_WORK)
    {
        Console.WriteLine("dead code");
    }
}
```

Poll twice?

Poll once and reuse the population?

Dead code – reason about collection properties and query/code interactions

CONDITIONAL BLENDING

```
var unemployed =  
    from person in people  
    where person.Employment==Employment.WORKED_FULL_TIME_YEAR  
    select person;  
  
foreach (var person in unemployed)  
{  
    if (person.Gender == Gender.FEMALE)  
    {  
        Console.WriteLine("unemployed female: {0}", person);  
    }  
}
```

```
var unemployedFemales =  
    from person in people  
    where  
        person.Employment==Employment.WORKED_FULL_TIME_YEAR  
        person.Gender == Gender.FEMALE  
    select person;  
  
foreach (var person in unemployedFemales)  
{  
    Console.WriteLine("unemployed female: {0}", person);  
}
```

CROSS-CROWD SAMPLING STRATEGIES

Having more than one crowd can help us assess the issue of **ignorability**: does the fact that we are **polling online** affect the outcome?

Test respective biases brought about by **different** crowds

Provide answers for different population **mixes**: 70% Indian, 30% American

Specific optimizations

- US -> India
- US and India, blended
- US during the day, India at night
- Use India to find missing population segments

MORE OPTIMIZATIONS

Can we cache answers to common questions and reuse them across surveys? Is that always the correct thing to do?

Can we determine questions that are time-sensitive?

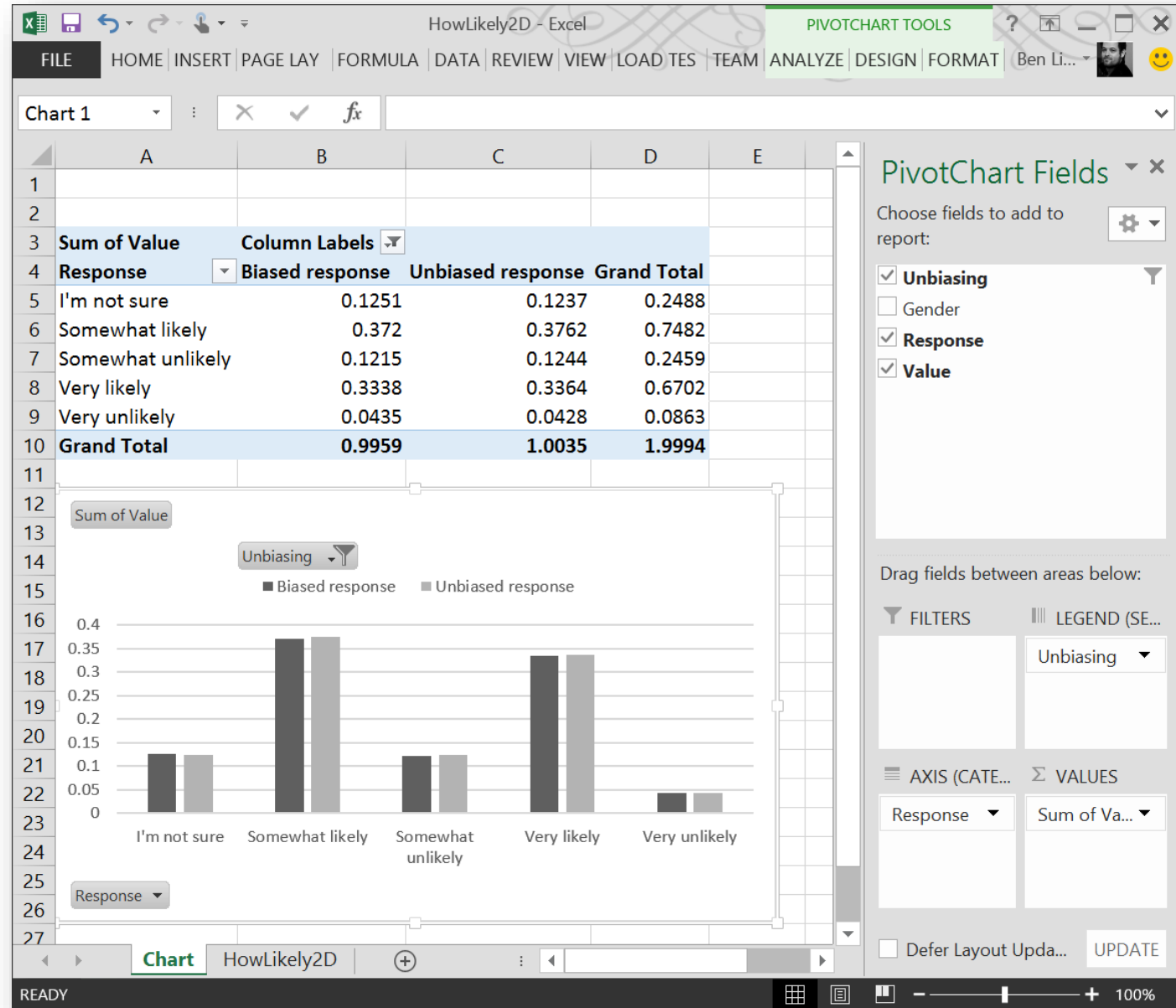
Can we decide if keeping answers to specific questions may lead to privacy violations.

Can we determine questions suitable for different crowds?

Can we determine if someone is likely to cheat or compensate for the possibility of cheating?

EXPLORING THE DATA

Age	Gender	Race	Education	Income	Employment	Relationship	Zip	Answer 9	Answer 10	Answer 11
48	Female	Two or more races	Bachelor's	\$75,000 or more	Worked full time	Married, Spouse	98275	From time to time	Definitely not at all	Not at all
24	Female	Black or African American	Bachelor's	\$5,000 to \$9,999	Worked full time	Never married	60638	Not at all	Only a little	Not at all
22	Male	White alone	High school or less	\$15,000 to \$24,999	Worked full time	Married, Spouse	81003	From time to time	Hardly at all	A little
31	Male	White alone	Bachelor's	\$35,000 to \$49,999	Worked full time	Married, Spouse	14226	A lot of times	Not quiet	Yes, but not a lot
18	Male	White alone	High school or less	\$1 to \$4,999	Did not work full time	Never married	77388	From time to time	Definitely not at all	A little
53	Male	White alone	Some college	\$75,000 or more	Worked full time	Married, Spouse	90606	A lot of times	Not quiet	A little
28	Female	Asian alone	Bachelor's	\$50,000 to \$74,999	Worked full time	Never married	98056	From time to time	Not quiet	Yes, but not a lot
50	Male	White alone	High school or less	\$75,000 or more	Worked full time	Married, Spouse	29730	From time to time	Definitely not at all	Not at all
22	Female	White alone	Bachelor's	\$1 to \$4,999	Worked full time	Never married	6612	From time to time	Not quiet	Yes, but not a lot
28	Male	White alone	Bachelor's	\$10,000 to \$14,999	Worked full time	Never married	95776	A lot of times	Hardly at all	Very likely
33	Female	White alone	Bachelor's	No income	Did not work full time	Married, Spouse	89012	From time to time	Not quiet	A little
30	Female	Two or more races	Some college	\$1 to \$4,999	Did not work full time	Never married	92869	From time to time	Not quiet	Not at all



INTERFACES

```
1 using ...
4
5 namespace Microsoft.Research.RiSE.InterPoll
6 {
7     public partial class Runner
8     {
9         [TestMethod]
10        public void EmploymentSurvey() {
11            }
12        }
13    }
14 }
```

CONCLUSIONS

InterPoll: a system for large-scale crowd-sourced polling

Geared toward

- Developers who want to incorporate human data into their applications
- But also social scientists
- Marketing professionals
- Campaign pollsters

Have explored **power analysis**
and are doing experiments
on **unbiasing**
and various **optimizations**