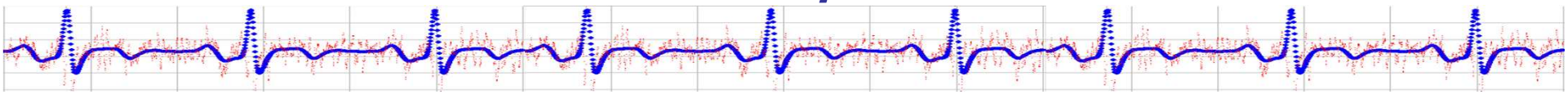
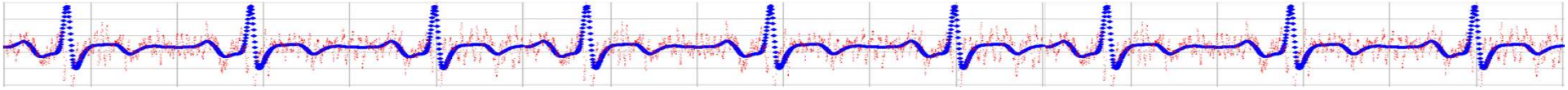
Empirical Research Methods in Information Science

IS 4800 / CS6350



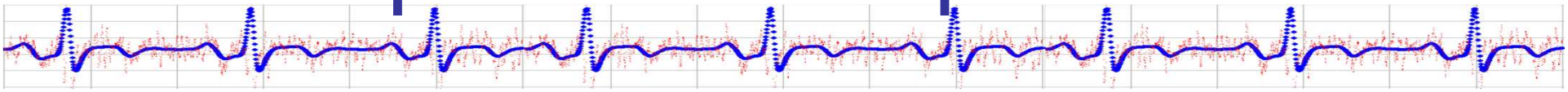
Lecture 11

Outline



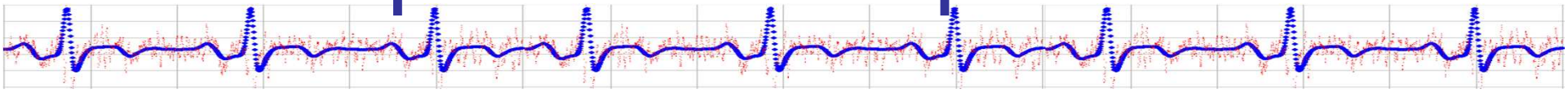
- Reading assessment
- Homework I4 questions
- Surveys
- Sampling
- Chi-square

Homework I4: Design a composite self-report measure



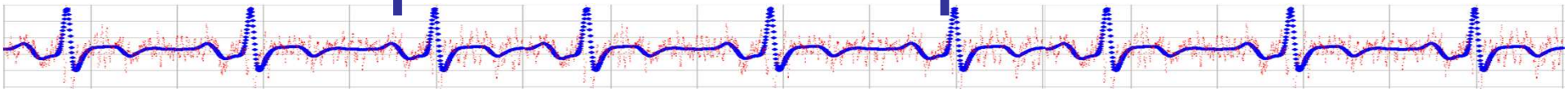
- Design a new composite self-report measure (e.g. "homework procrastination") … but your own idea
- Assume it only has one factor, but use at least five scale items
- Incorporate information from at least one literature reference

Homework I4: Design a composite self-report measure



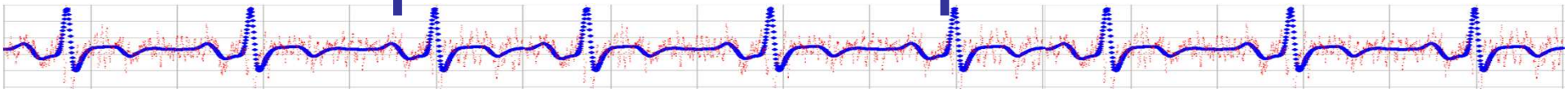
- Assess the face and content validity of your measure and work through a bivariate analysis of your items
- Implement questionnaire on surveymonkey.com or Google forms

Homework I4: Design a composite self-report measure



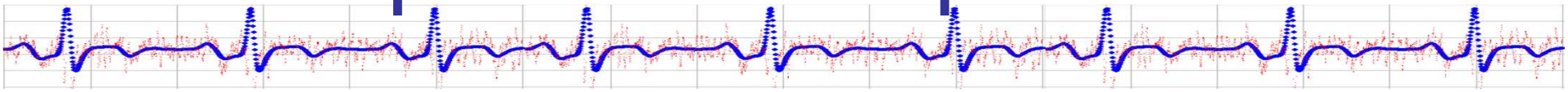
- Decide on one method for assessing validity (besides face & content) for your measure that you can also assess in a self-report questionnaire. This should be an additional question (or an additional previously validated composite measure) on your survey and should provide a numeric measure

Homework I4: Design a composite self-report measure



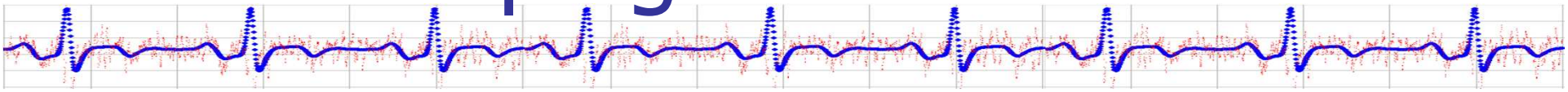
- Post your questionnaire on Piazza
- You are obligated to reply to any questionnaires posted within 48 h!
- Compute the reliability (internal consistency) of your measure using Python
- Compute descriptive statistics for your measure and any other items you may have included on the questionnaire

Homework I4: Design a composite self-report measure



- Assess the validity of your measure (you can do this qualitatively, e.g., using scatterplots)
- Document and submit all of the above
- You may work individually or in teams of two
- Due 2/20

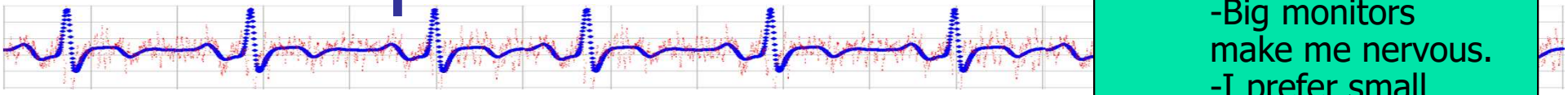
Developing a new measure



- Say you decide you need a new survey measure, "attitude towards large computer monitors" (ATLCM)
  - I like big monitors.
  - Big monitors make me nervous.
  - I prefer small monitors, even if they cost more.
  - *7-pt Likert scales*

- How would you validate this measure?

Overall process to de composite measure

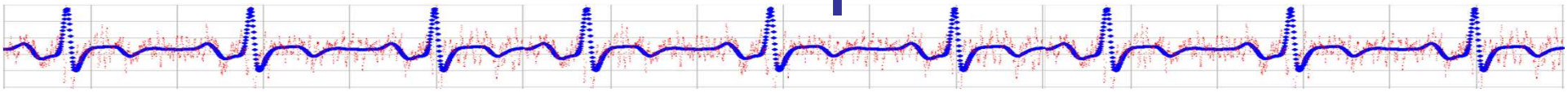


"Attitude towards large computer monitors" (ATLCM)
-I like big monitors.
-Big monitors make me nervous.
-I prefer small monitors, even if they cost more.
-7-pt Likert scales

- Identify factors
- Identify items
- Face and content validity for each item
- Check response variance for each item
  (Check floor/ceiling effects)
- Bi-variate analysis
- Test reliability
- Test validity

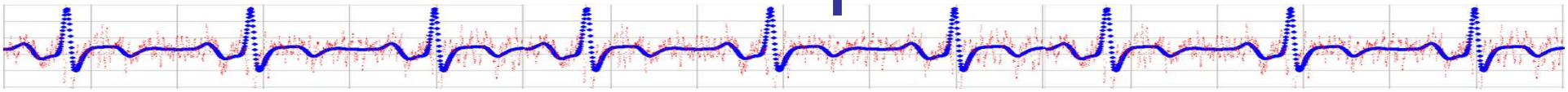
Good checklist for assignment!

Chronbach's alpha



- Measure of internal *consistency*
  - Not homogeneity
  - Not unidimensionality (assumes all items measure single dimension … test multi-dimensionality using factor analysis)
- Function of the number of test items and the average inter-correlation among the items
- Average of all possible split halves

Chronbach's alpha



- Be careful what you report!
  "Reliability of the subscale scores was…"
  (*Not* "Reliability of the subscale was…")

- May have been superseded by omega (available in R, looking for Python implementation)

Chronbach's alpha

$$\alpha = \frac{N \cdot \bar{c}}{\bar{v} + (N-1) \cdot \bar{c}}$$

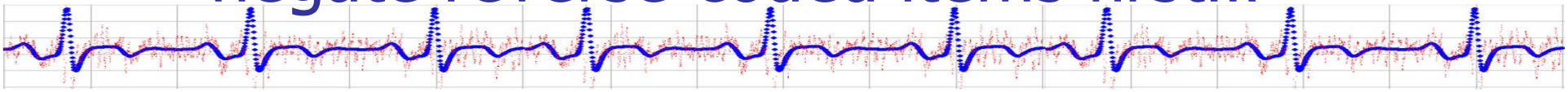
- N = number of items
- c-bar is the average inte
  covariance among the items
- v-bar equals the average variance

("Standardized alpha" based on correlations instead of covariance)

- Increase # items, increase alpha
- If average inter-item correlation is low, alpha will be low. As the average inter-item covariance increases, Cronbach's alpha increases as well (holding the number of items constant).

Cronbach's alpha

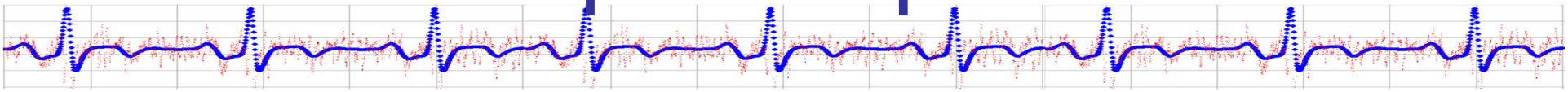
negate reverse-coded items first…



- Scores
  - .00 (no consistency) to 1.0 (perfect)
  - .70 (70% of variance reliable variance)
    Commonly cited as acceptable
    (Something important, might want .90+)

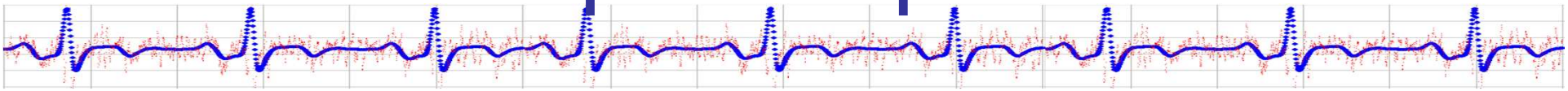
- See Python guide to compute…

Increasing the reliability of a composite questionnaire



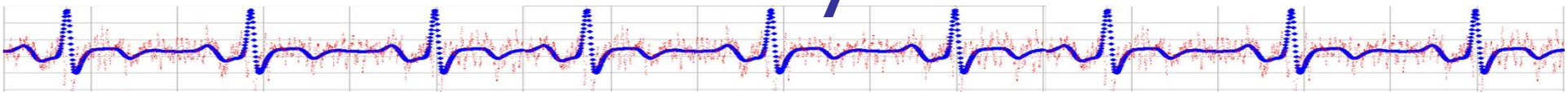
- Increase the number of items on your questionnaire
- Standardize the conditions under which the test is administered (e.g., timing procedures, lighting, ventilation, instructions)
- Make sure you score your questionnaire carefully, eliminating scoring errors

Increasing the reliability of a composite questionnaire

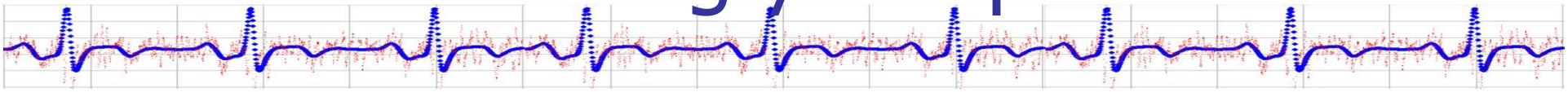


- Check to be sure the items on your questionnaire are clearly written and appropriate for those who will complete your questionnaire
- Assess reliability with each item dropped
  (In R, see:
  https://cran.r-project.org/web/packages/psych/psych.pdf)

Questionnaire administration & results analysis

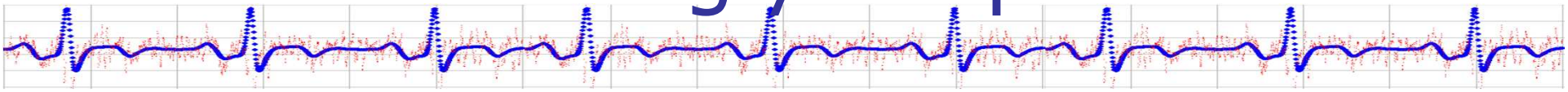


Administering your questionnaire



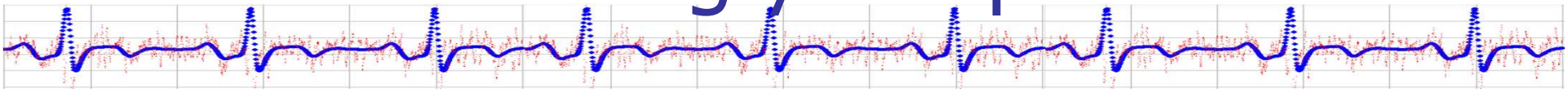
- Mail survey
  - A questionnaire is mailed directly to participants
  - Convenient
  - Nonresponse bias is a serious problem resulting in an unrepresentative sample

Administering your questionnaire



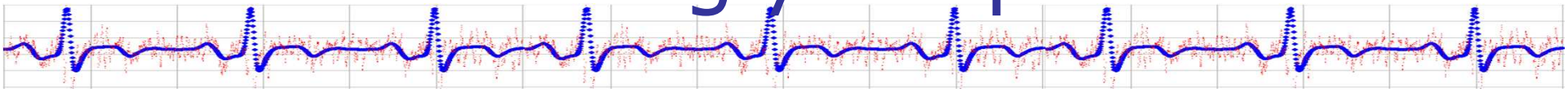
- Mail survey
- Internet survey
  - Distributed via e-mail or on a website
  - Large samples can be acquired quickly
  - Biased samples are possible because of uneven computer ownership across demographic groups
  - Check out surveymonkey.com or Qualtrics

Administering your questionnaire



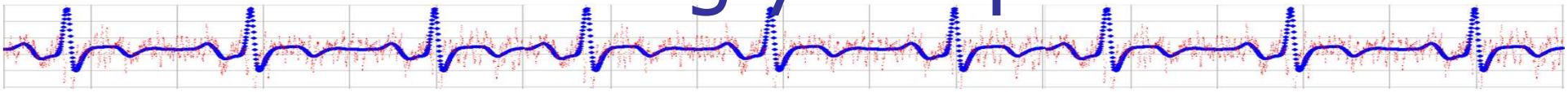
- Mail survey
- Internet survey
- Telephone survey
  - Participants are contacted by telephone and asked questions directly
  - Questions must be asked carefully
  - The plethora of "junk calls" may make participants suspicious

Administering your questionnaire



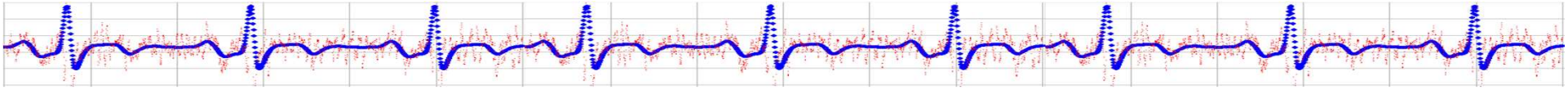
- Mail survey
- Internet survey
- Telephone survey
- Group administration
  - A questionnaire is distributed to a group of participants at once (e.g., a class)
  - Completed by participants simultaneously
  - Ensuring anonymity may be a problem

Administering your questionnaire



- Mail survey
- Internet survey
- Telephone survey
- Group administration
- Interview
  - Participants asked questions face-to-face
  - Characteristics/behavior of the interviewer may affect the participants' responses

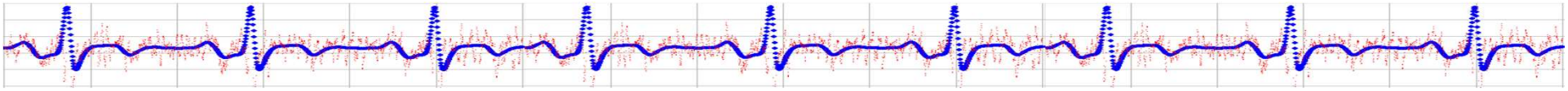
Mechanical Turk



Invented by Hungarian engineer and inventor Baron Wolfgang von Kempelen in 1769.

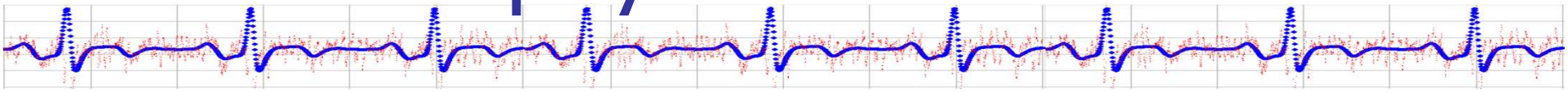
Photo a reconstructed version by John Gaughan.

Mechanical Turk



- Amazon Mechanical Turk: crowdsourcing tool
- People perform small tasks for micropayments
- Developed by Peter Cohen at Amazon for its internal use (find duplicates in webpages in 2005)
- Boon for certain type of research projects

MTurk payment

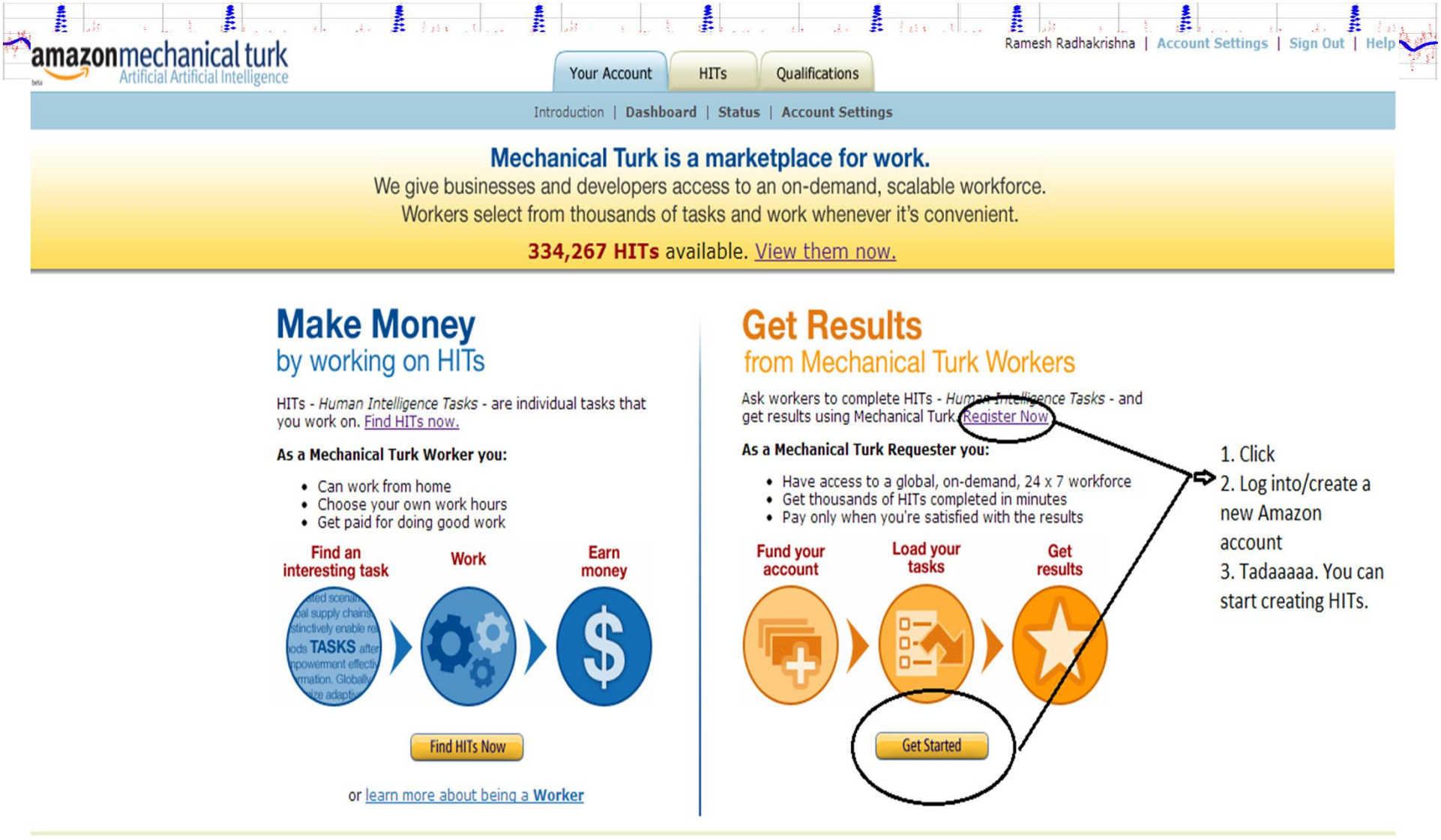


2017 paper:

- Median hourly wage: ~$2/h
- Only 4% earned more than $7.25/h
- Average requester pays more thank $11/h
- But lower-paying requesters post more work

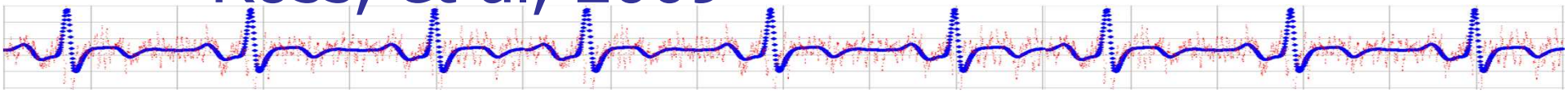
Hara et al. “A Data-Driven Analysis of Worker’s Earnings on Amazon Mechanical Turk,” CHI 2018.

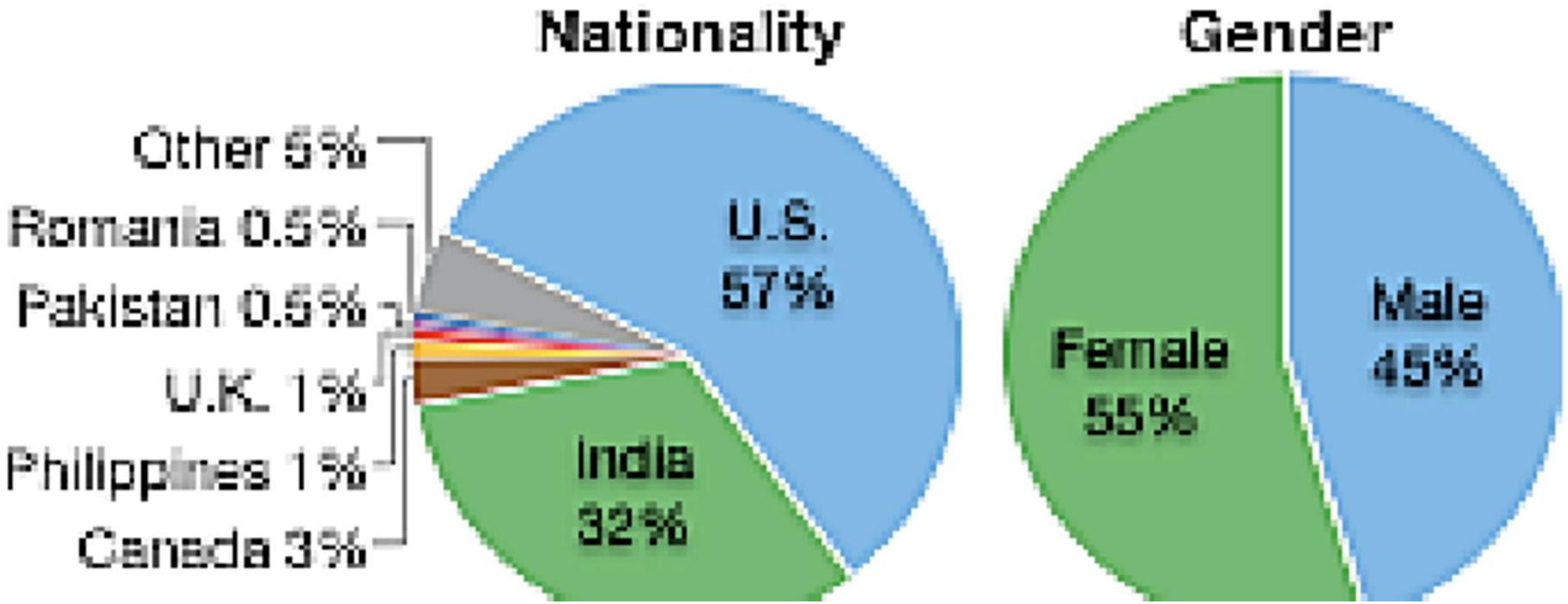
Mechanical Turk



Turker demographics

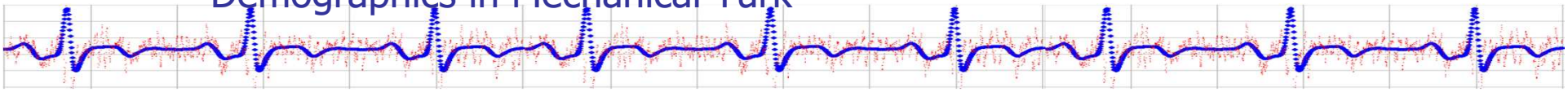
Ross, et al, 2009

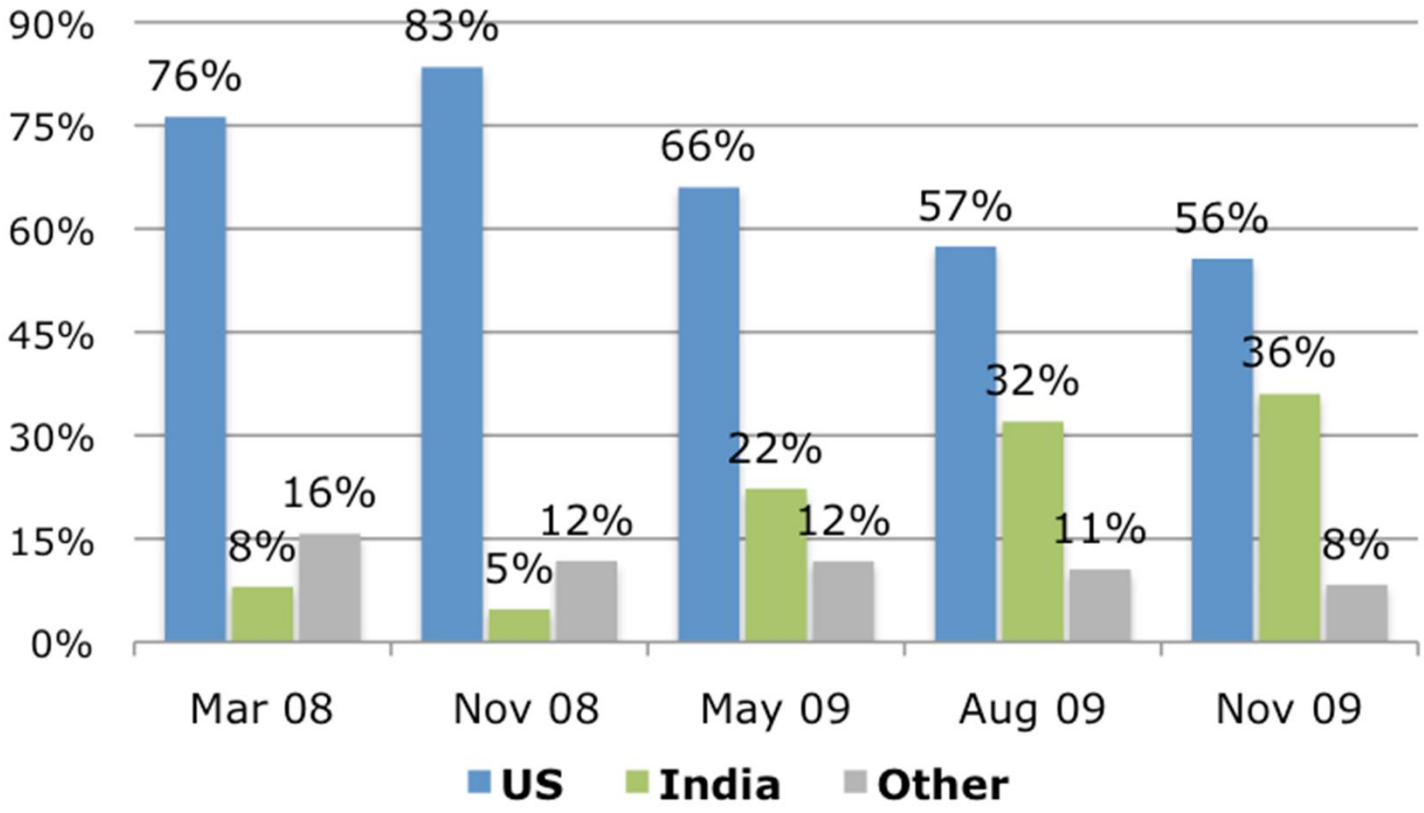




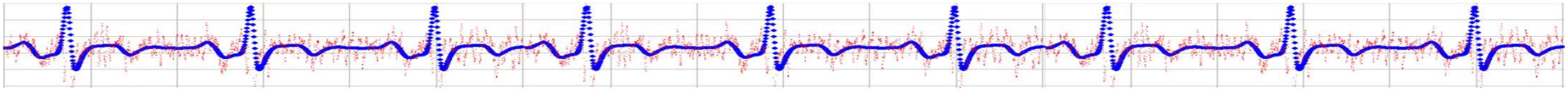
Turk: Demographics

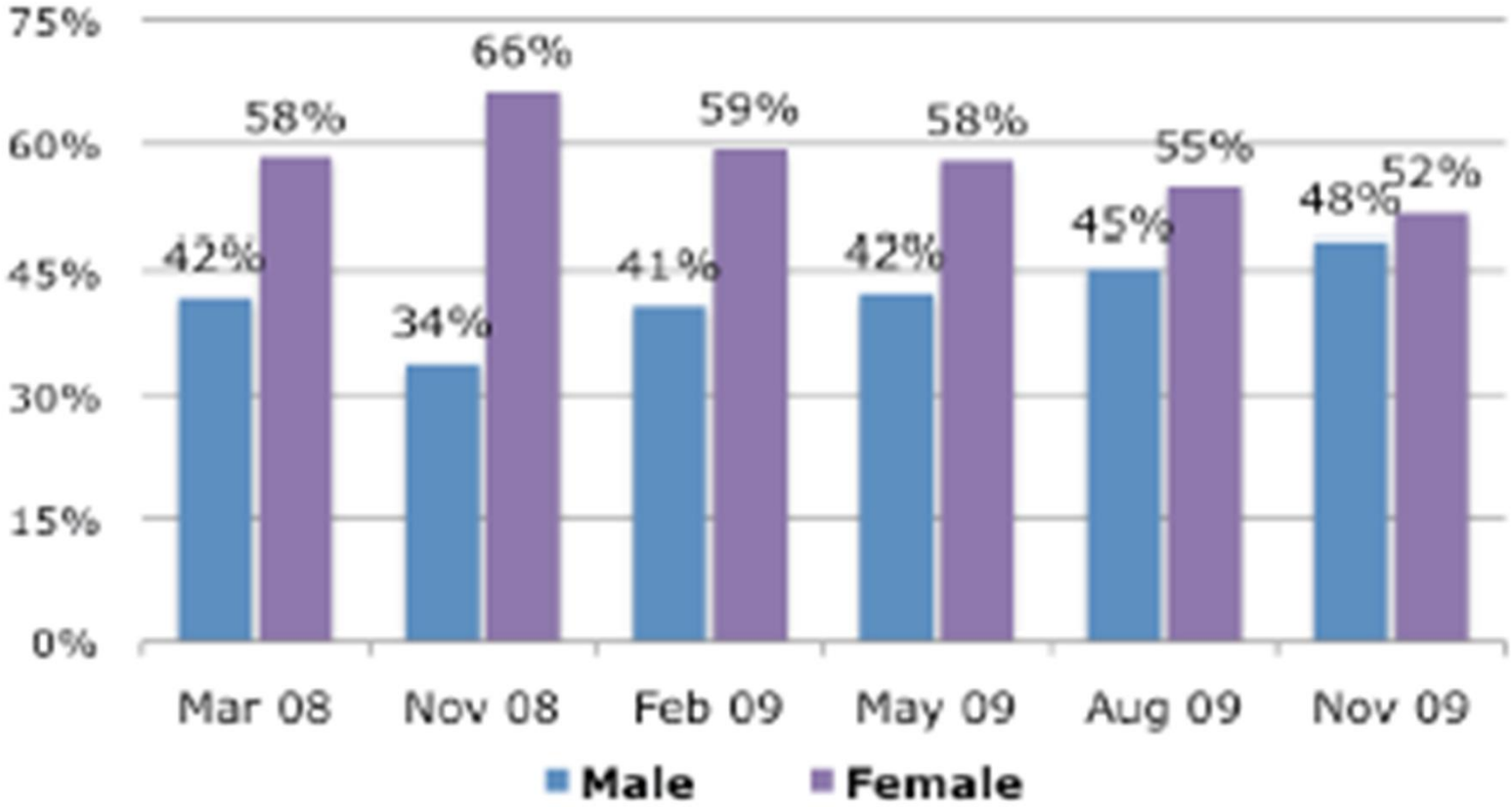
Ross, et al, CHI 2010, Who are the Crowdworkers? Shifting Demographics in Mechanical Turk



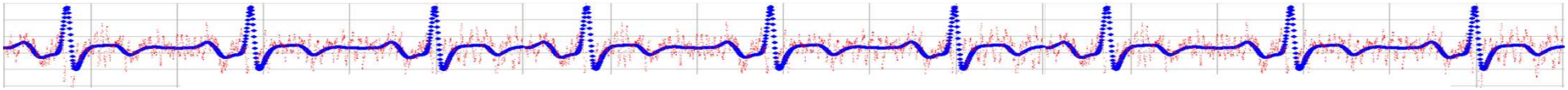


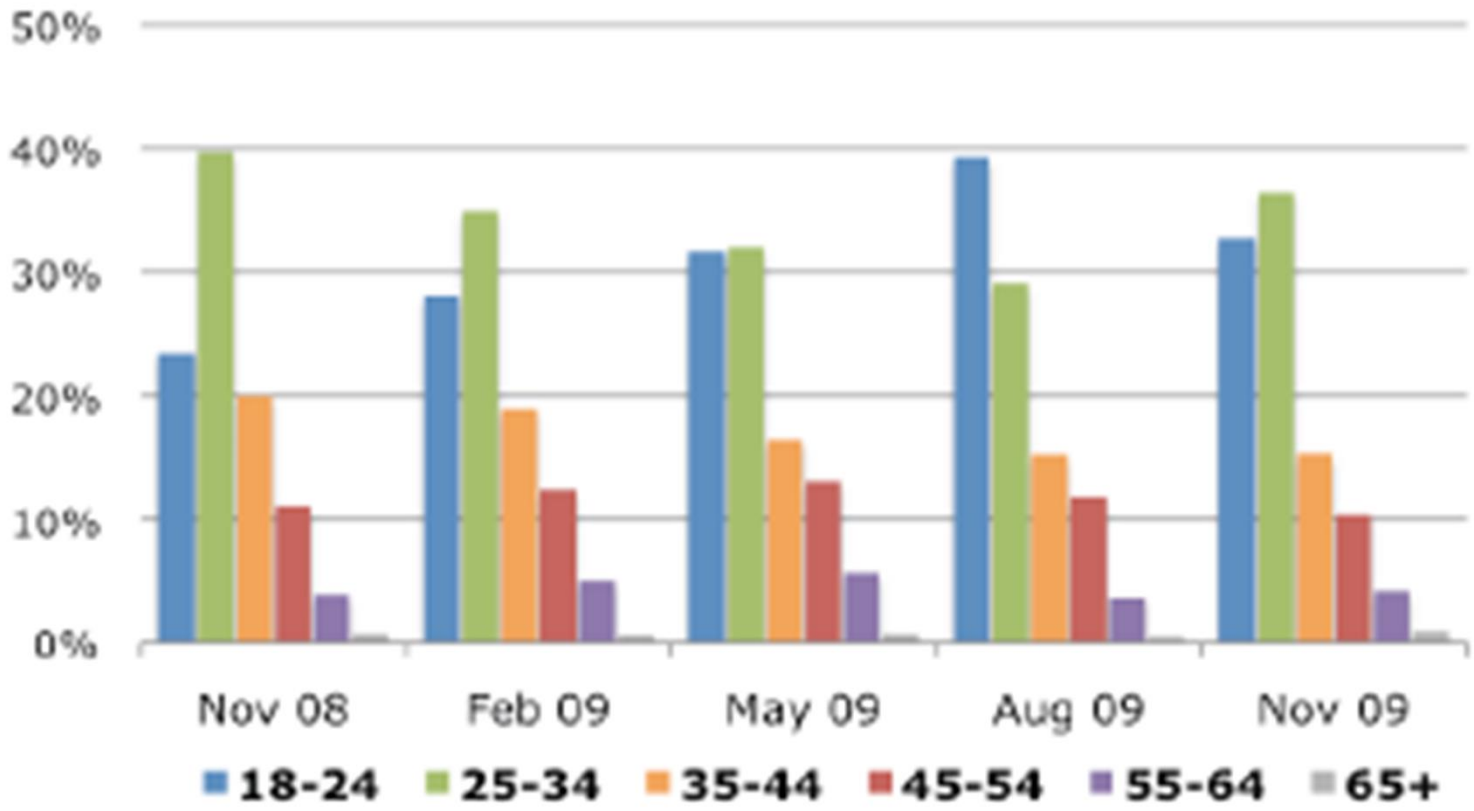
Turk: Demographics

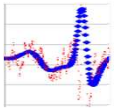


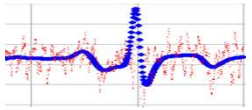


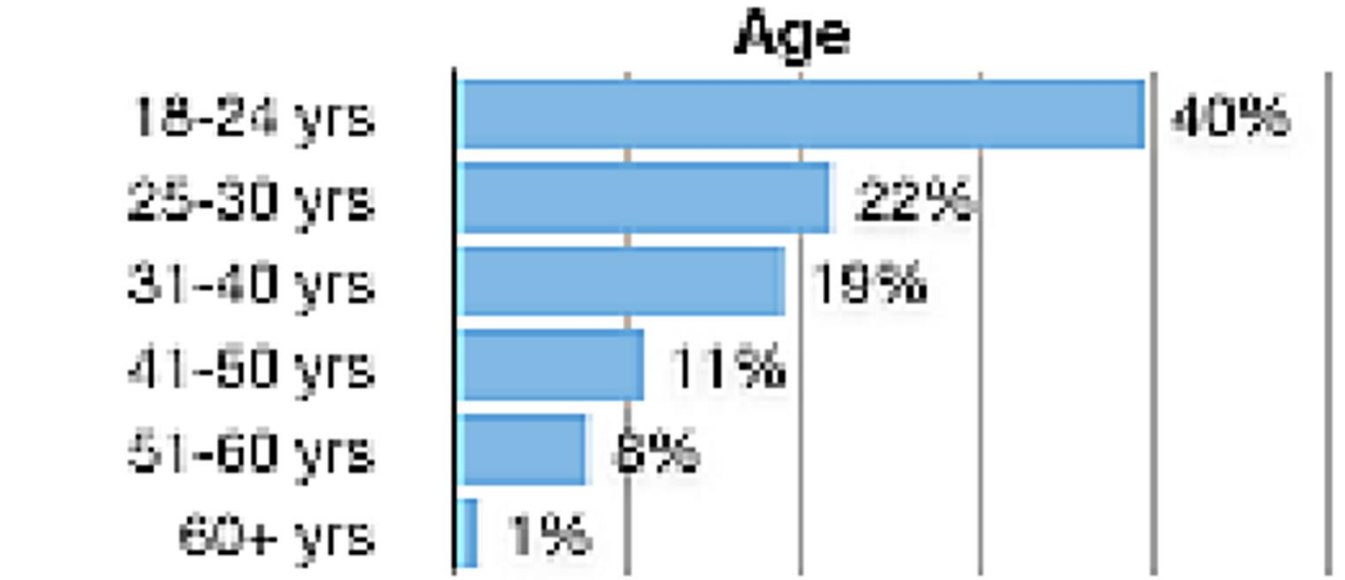
Turk: Demographics

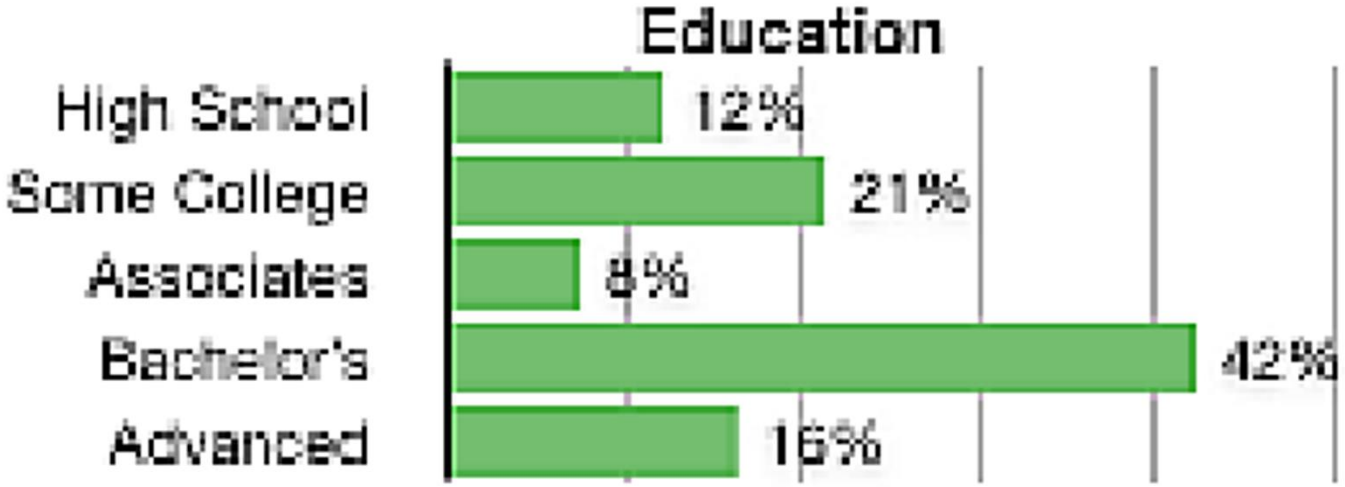


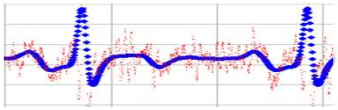


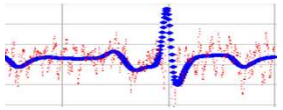


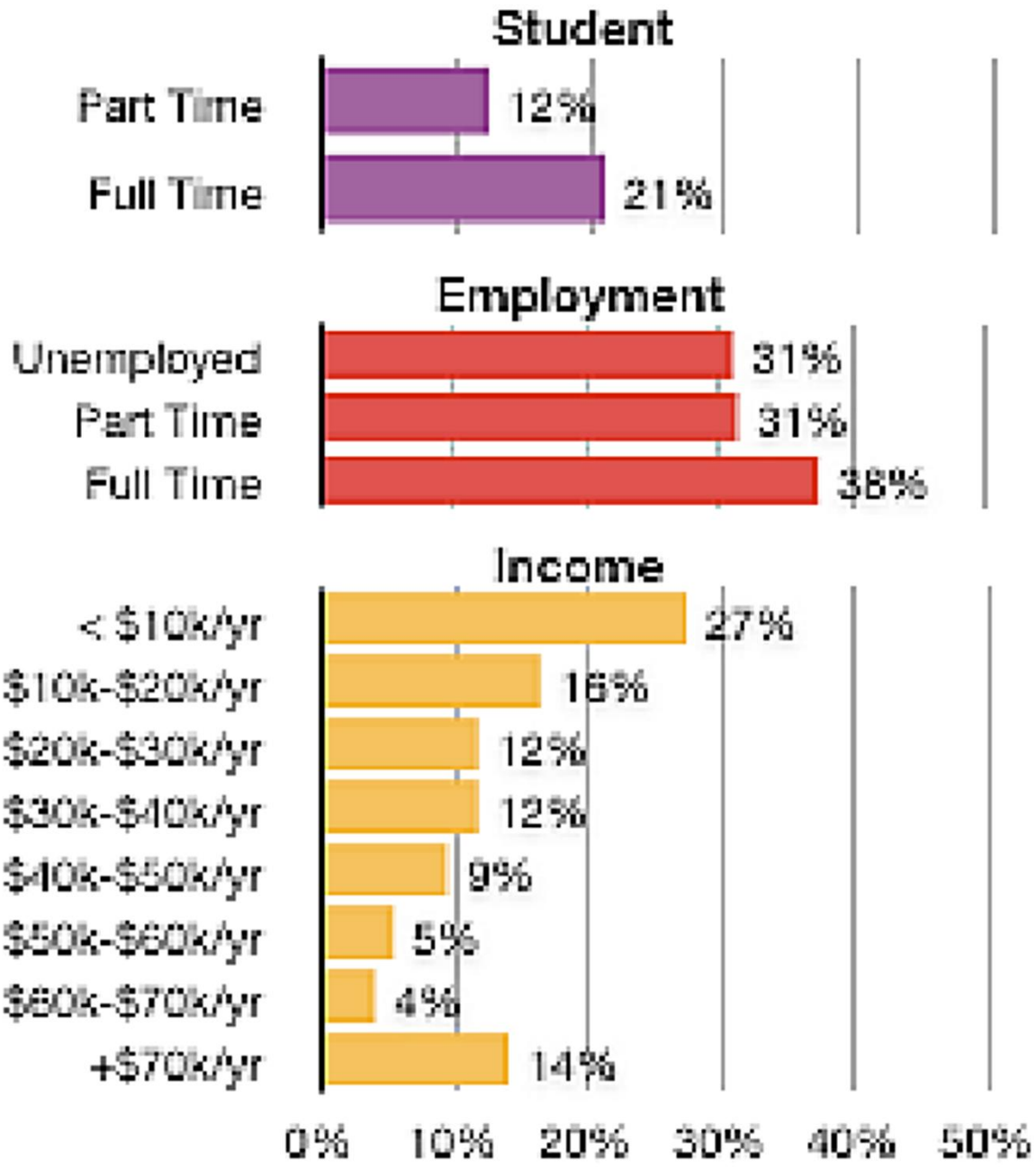



Age
18-24 yrs
40%
25-30 yrs
22%
31-40 yrs
19%
41-50 yrs
11%
51-60 yrs
6%
60+ yrs
1%

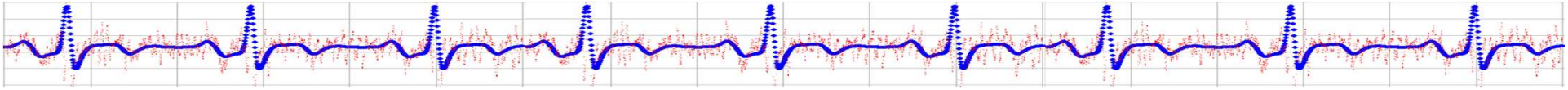

Education
High School
12%
Some College
21%
Associates
8%
Bachelor's
42%
Advanced
16%






Student
Part Time
12%
Full Time
21%
Employment
Unemployed
31%
Part Time
31%
Full Time
38%
Income
< $10k/yr
27%
$10k-$20k/yr
16%
$20k-$30k/yr
12%
$30k-$40k/yr
12%
$40k-$50k/yr
9%
$50k-$60k/yr
5%
$60k-$70k/yr
4%
+$70k/yr
14%
0%
10%
20%
30%
40%
50%

Or not!



A Medium Corporation [US] | https://medium.com/@silberman/stop-citing-ross-et-al-2010-wh...

Become a member Sign in Get started

Stop citing Ross et al. 2010, "Who are the crowdworkers?"

Six Silberman Follow
Mar 16, 2015 · 3 min read

M. Six Silberman, Kristy Milland, Rochelle LaPlante, Joel Ross, and Lilly Irani, with approval of Andrew Zaldivar and Bill Tomlinson

This paper is out of date and should no longer be cited.

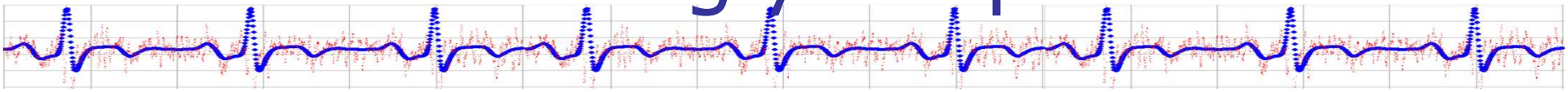
The average wage statistic presented in the paper is misleading, as are many other statistics such as those on the nationalities of workers.

While the statistics may have been accurate in 2009, since then:

Market demographics have changed

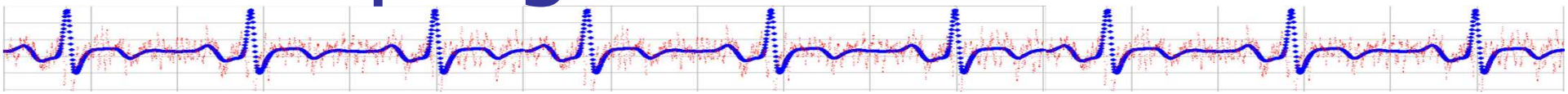
Market demographics have changed significantly due to system design and

Administering your questionnaire

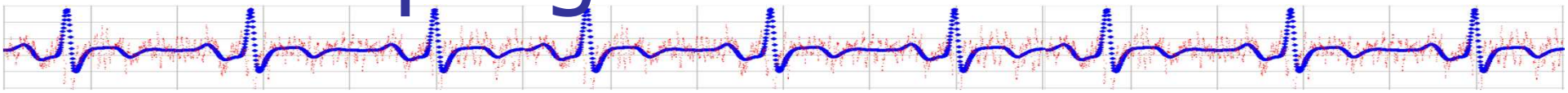


In general, personal techniques (interview, phone) provide higher response rates, but are more expensive and may suffer from bias problems

Sampling



Sampling

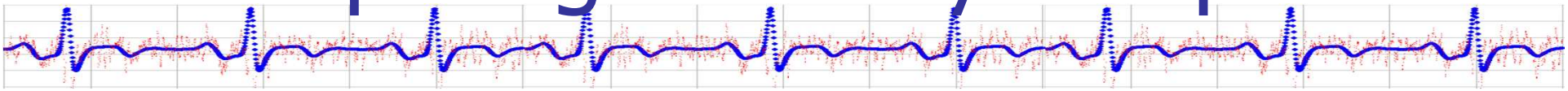


- Sometimes you really can measure the entire population (e.g., workgroup, company), but this is rare…

- "Convenience sample"

  Cases are selected only on the basis of feasibility or ease of data collection

Acquiring a survey sample

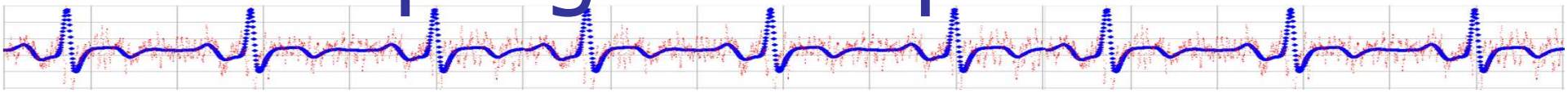


- Obtain a *representative sample*

  Sample closely matches population characteristics

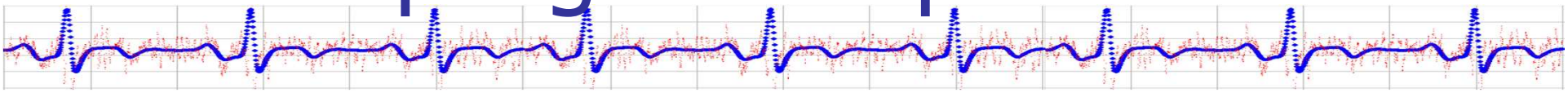
- A *biased sample* occurs when your sample characteristics don't match population characteristics
  - Produce misleading or inaccurate results
  - Usually stem from inadequate sampling procedures

Sampling techniques



- Simple random sampling
  - Randomly select a sample from the population
  - *Random digit dialing* is a variant used with telephone surveys
  - Reduces systematic bias, but does not guarantee a representative sample
    - Some segments of the population may be over- or underrepresented

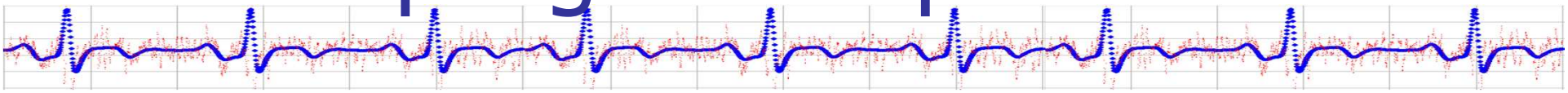
Sampling techniques



- Simple random sampling
- Systematic sampling
  - Every kth element is sampled after a randomly selected starting point

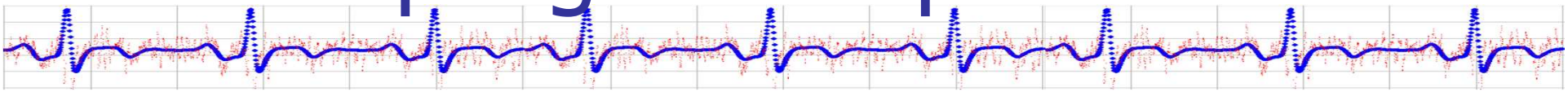
    E.g., sample every 5th name in the telephone book after a random page/start point
  - Empirically equivalent to random sampling, but not always
  - Easier than random sampling

Sampling techniques

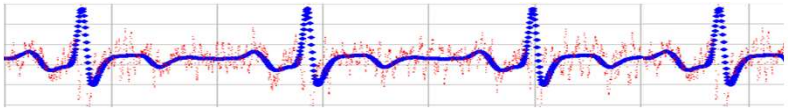


- Simple random sampling
- Systematic sampling
- Stratified sampling
  - Used to obtain a representative sample
  - Population divided into (demographic) strata
  - Focus also on variables that are related to other variables of interest in your study (e.g., relationship between age and computer literacy)
  - A random sample of a fixed size is drawn from each stratum
  - May still lead to over- or underrepresentation of certain segments of the population

Sampling techniques

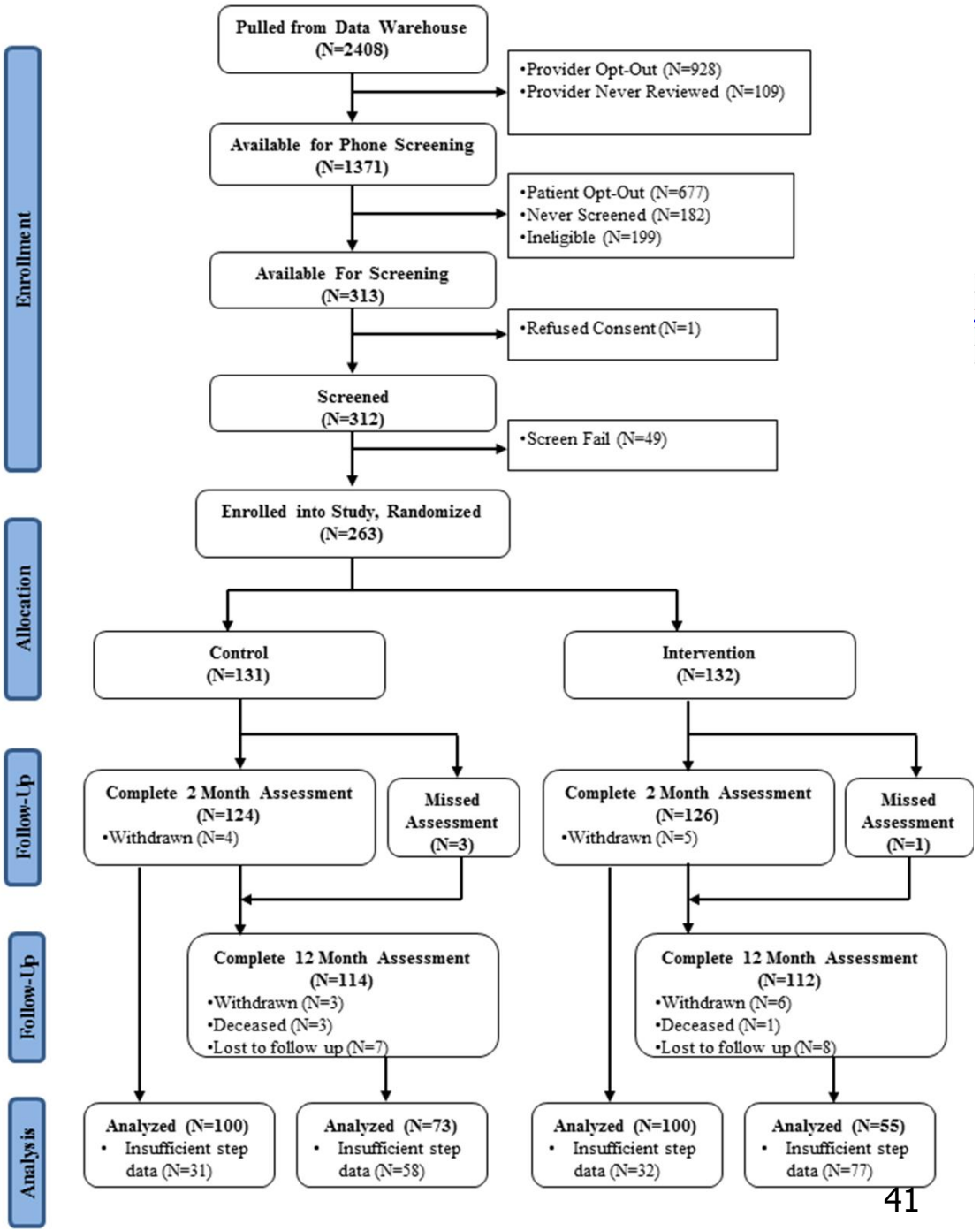


- Simple random sampling
- Systematic sampling
- Stratified sampling
- Proportionate sampling
  - Same as stratified sampling except that the proportions of different groups in the population are reflected in the samples from the strata



Example

Stratified Sampling

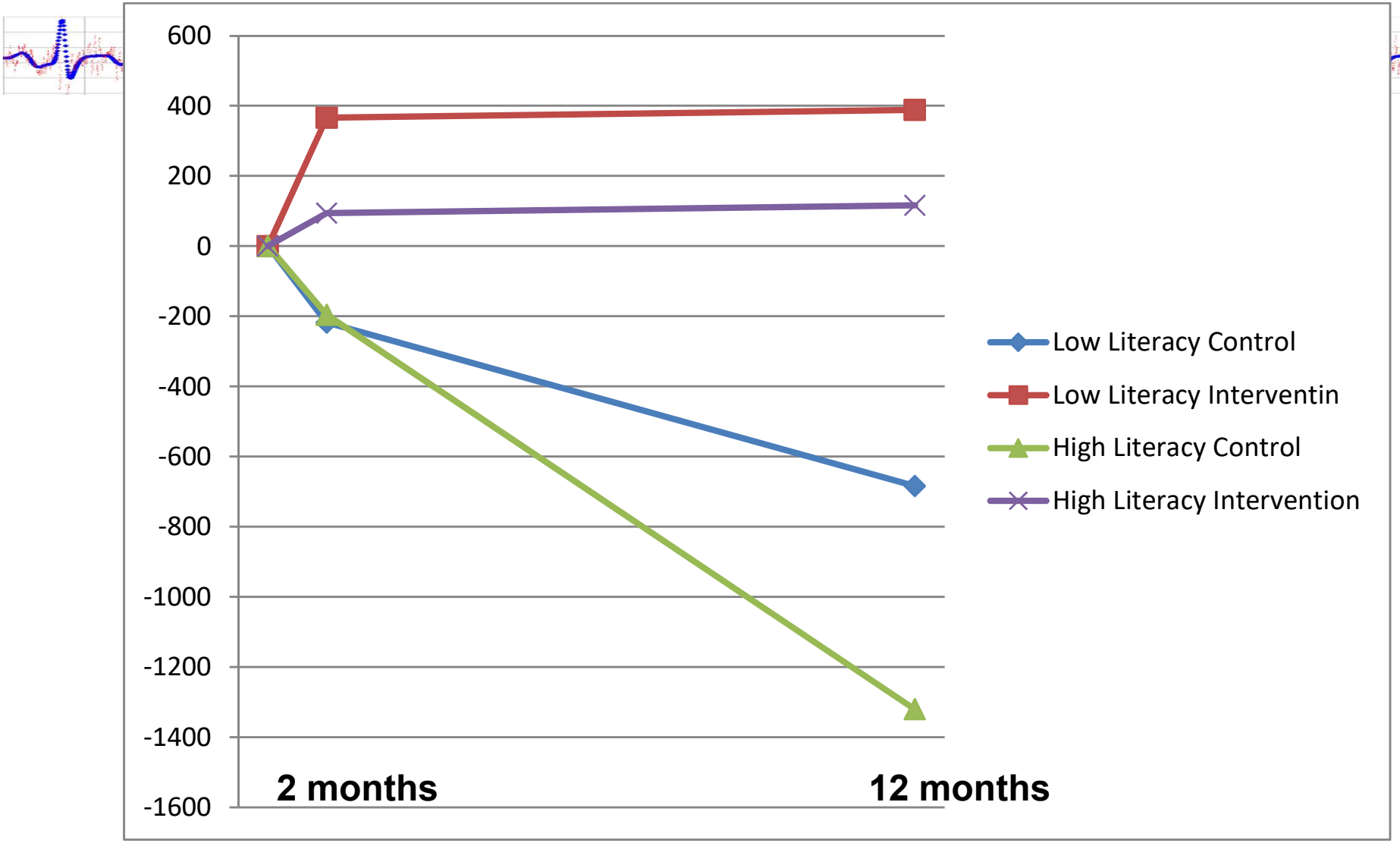


Low Literacy Geriatrics Pts

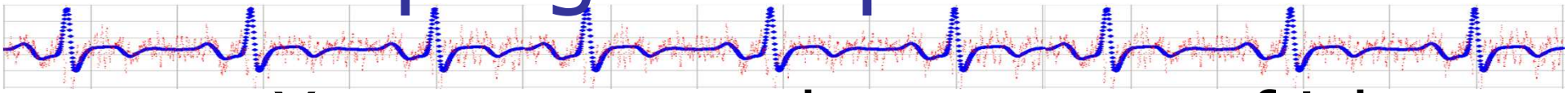
NIA R01, N=263, 55+

National Institute on Aging
Intramural Research Program

Low Literacy Geriatrics Pts

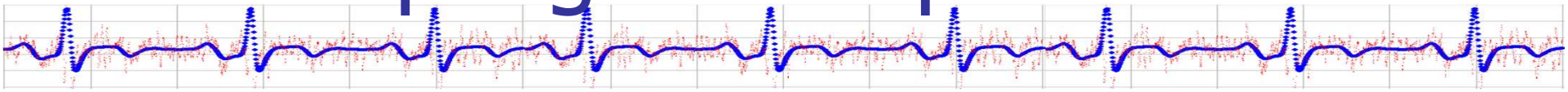


Sampling example:



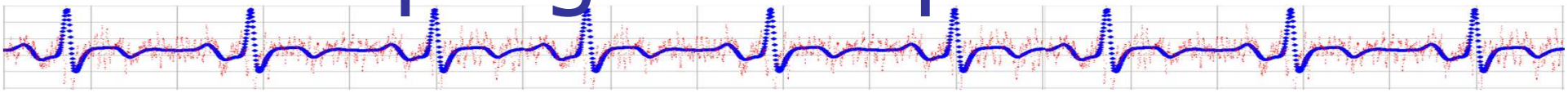
- You want to conduct a survey of job satisfaction of all employees but can only afford to contact 100 of them
- Personnel breakdown:
  - 50% Engineering
  - 25% Sales & Marketing
  - 15% Admin
  - 10% Management
- Examples of
  - Stratified sampling?
  - Proportionate sampling?

Sampling techniques



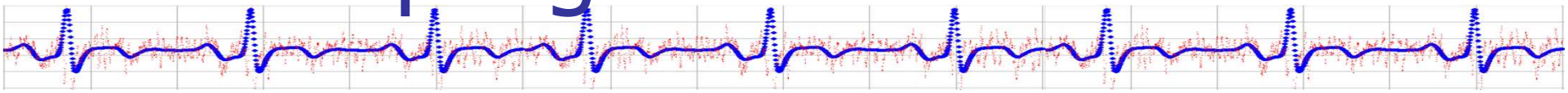
- Cluster Sampling
  - Used when populations are very large
  - The unit of sampling is a group (e.g., a class in a school) rather than individuals
  - Groups are randomly sampled from the population (e.g., ten classes from a particular school)

Sampling techniques



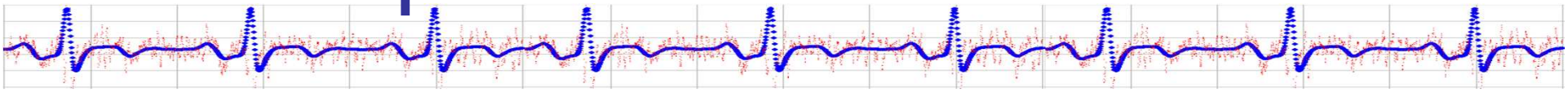
- Cluster sampling
- Multistage sampling
  - Variant of cluster sampling
  - First, identify large clusters (e.g., school districts) and randomly sample from that population
  - Second, sample individuals from randomly selected clusters
  - Can be used along with stratified sampling to ensure a representative sample
  - Note: Multilevel, hierarchical statistical analysis can tease apart differences: individual vs. cluster

Sampling



Most statistics assume a random sample!

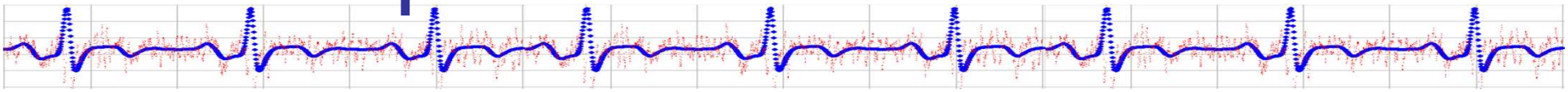
Sample size



- In all empirical research, you should motivate your *sample size*

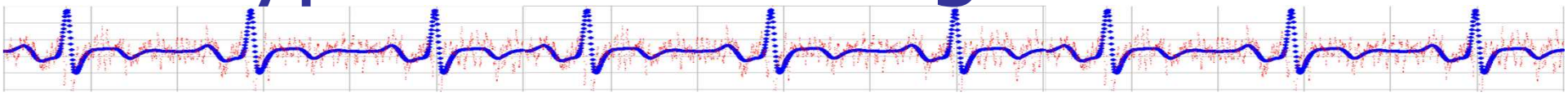
- B&A Ch 9 provide formula for estimating sample size for binomial descriptive studies.
  - For binomial (two category) measures
  - Based on
    - Amount of acceptable *sampling error*
    - Expected magnitude of population proportions

Sample size

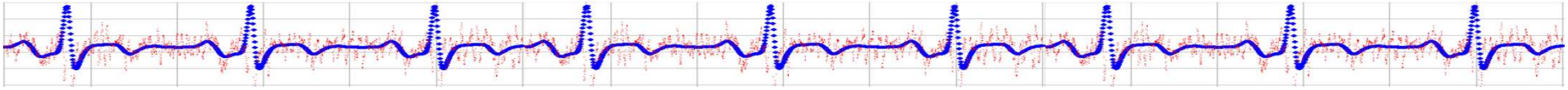


- You should try to select an *economic sample*
  - Includes enough respondents to ensure a valid survey and no more
- Two factors are taken into account when determining necessary sample size
  - Amount of acceptable *sampling error*
  - Expected magnitude of population proportions
- Formula to calculate sample size using the above parameters

Hypothesis Testing Preview

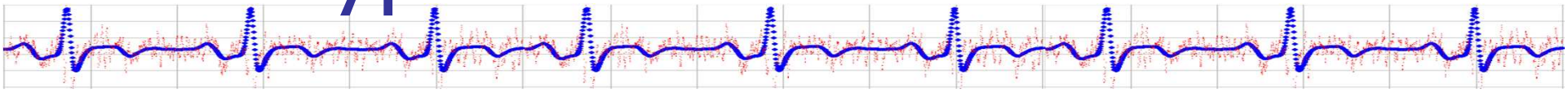


Lecture



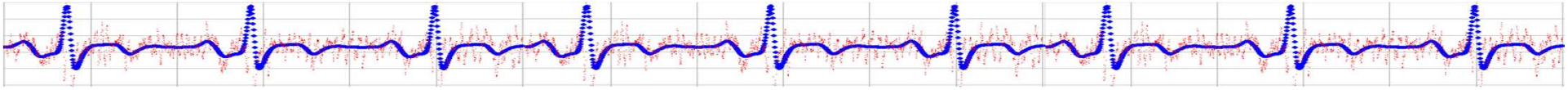
- How would you do hypothesis testing?
- Kinds of errors?
- Which are worse?

A typical scenario



- Between-subject design
- Let every subject try both Wizziword & Word
- Measure performance
- Research Hypothesis:
  - Wizziword is better

What if …

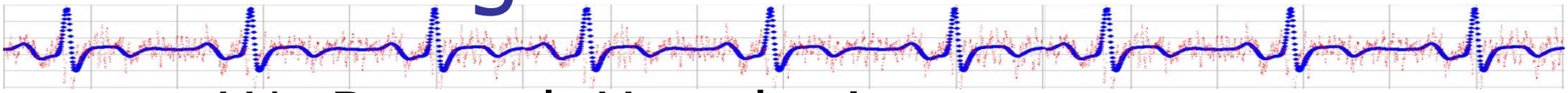


- You can test every subject in your population, AND
- There is no measurement error?
- Nothing else that could cause "error variance"

→

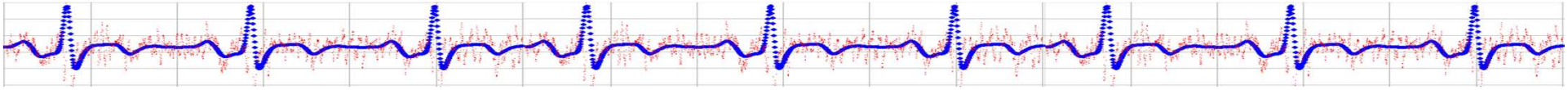
- Compute descriptives for two treatments
- If Wizziword perf > Word perf conclude H1 is supported
- No uncertainty, No 'p' value

Basic Process of Hypothesis Testing



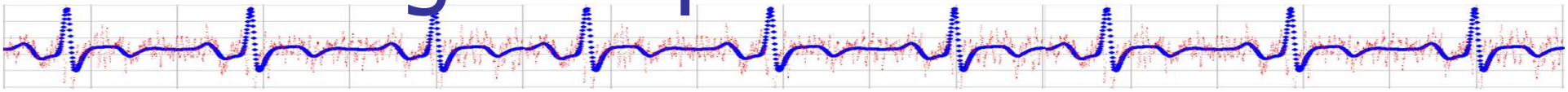
- H1: Research Hypothesis:
  - Population 1 is different than Population 2
- H0: Null Hypothesis:
  - No difference between Pop 1 and Pop 2
  - *The difference is "null"*
- Compute *p(observed difference|H0)*
  - 'p' = probability observed difference is due to random variation
- If p<*threshold* then reject H0 => accept H1
  - p typically set to 0.05 for most work
  - p is called the "level of significance"

Other ways of thinking about this…



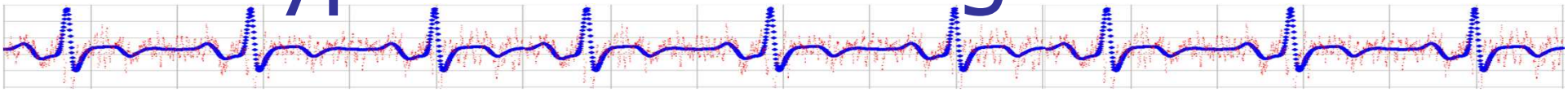
- "Innocent until proven guilty."
- How surprising would this result be if there really were no difference?
- Why do things this way???

The grand plan



- X^2 tests
  - For nominal measures
  - Can apply to a single measure
- Correlation tests
  - For two numeric measures
- t-test for independent means
  - For categorical IV, numeric DV

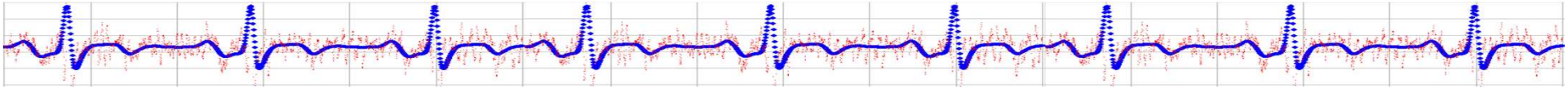
Hypothesis Testing



- How can we tell if our hunches about the world are right?
  - *Step 3 of the scientific method*

- How can we do this in the face of sampling error, measurement error, and extraneous variables?

A common scenario



IV **DV**

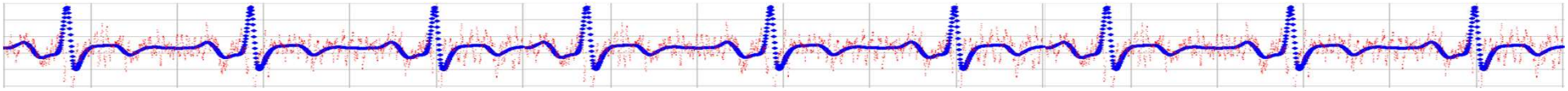
Software A vs. Software B

Efficiency (time to complete)

Sampling Error
Measurement Error
Covariates
Demand Effects
etc.

- **We measure difference A-B**
- **Is this real? Or just chance?**
- **We have very good models of randomness (noise).**
- **We can compute**
  p(observed difference A-B is due only to chance variation)
- **When should we conclude A-B is real (H1)?**
- **What can we conclude otherwise (H0)?**

What if …



- You can test every subject in your population, AND
- There is no measurement error? AND
- There are no covariates or other noise?

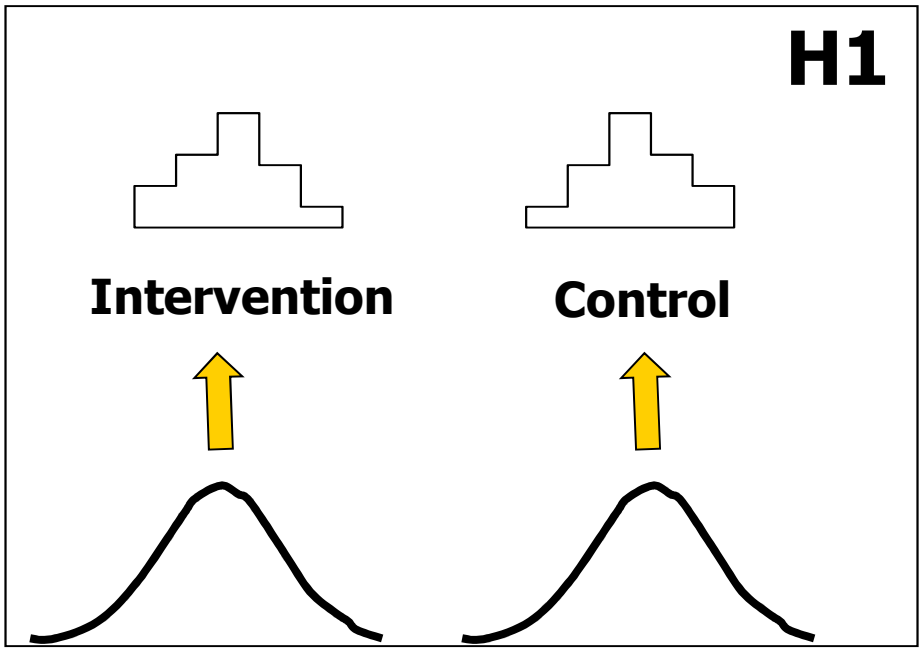
→

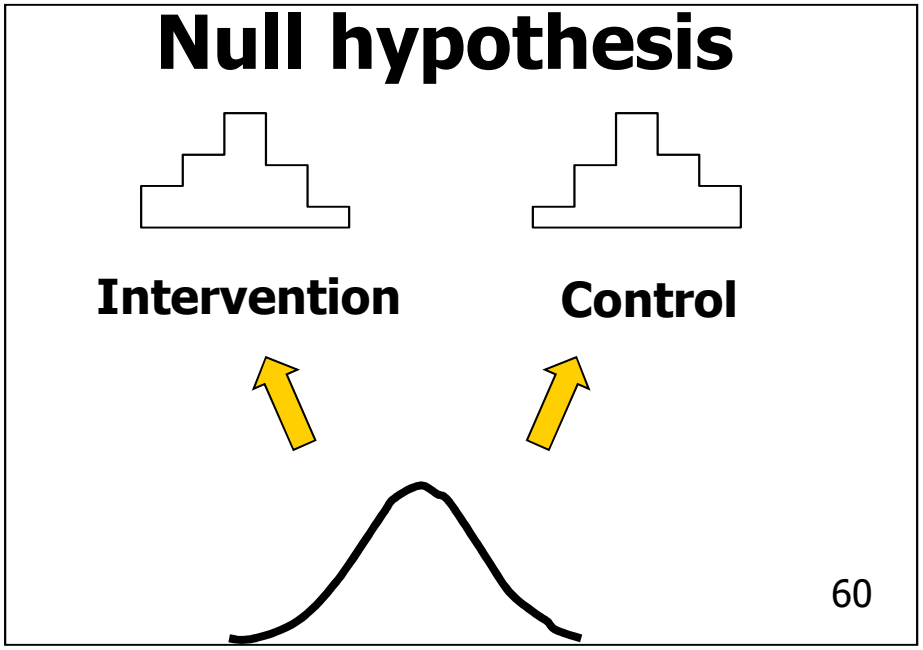
- Compute descriptives for two treatments
- If Software A Efficiency > Software B Efficiency conclude H1 is supported
- No uncertainty, p=0.0

Intervention

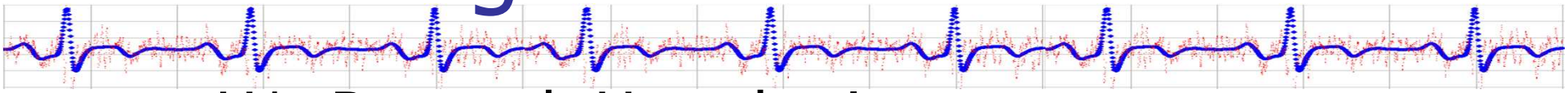
Control

The big question: which is correct?



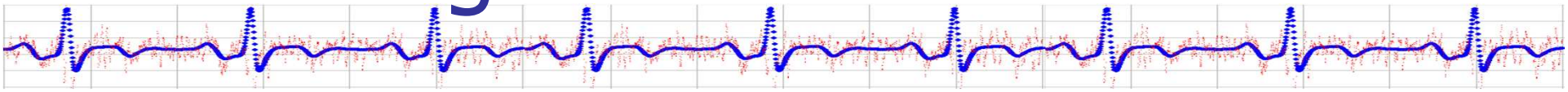


Basic Process of Hypothesis Testing



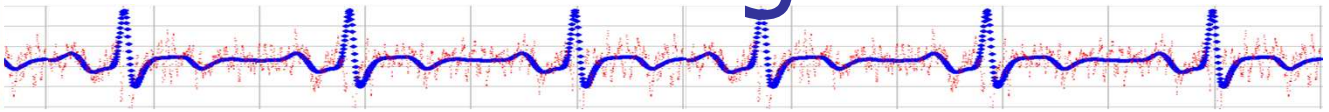
- H1: Research Hypothesis:
  - Population 1 is different than Population 2 on some measure
- H0: Null Hypothesis:
  - No difference between Pop 1 and Pop 2
  - *The difference is "null"*
- Compute *p(observed difference|H0)*
  - 'p' = probability observed difference is due to random variation
- If p<*threshold* then reject H0 => accept H1
  - p typically set to 0.05 for most work
  - p is called the "level of significance"

The plan for hypothesis testing



- X^2 tests
  - For 1 or 2 nominal measures
- Correlation tests
  - For two numeric measures
- t-test for independent means
  - For categorical IV, numeric DV

Survey Feb 5, 2013 Guns in Congress



- How many members of congress own guns?
- Dataset:

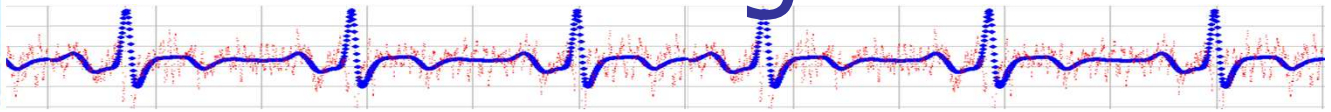
Title Member Party State GunOwner NRAGrade

- 531 seated members
- 165 Yes, 147 No, 219 “no comment”

Survey Feb 5, 2013 Guns in Congress

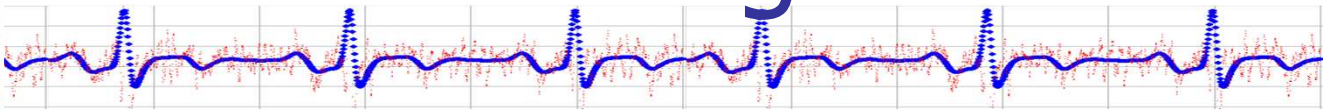
- Question: Is congressional gun ownership greater than US national gun ownership?
  - 32% of households have guns
    - 2010 data, 2011 study University of Chicago's National Opinion Research Center
    - Note: rate of individual ownership is less
  - Congress: 165 Yes, 147 No

Survey Feb 5, 2013 Guns in Congress



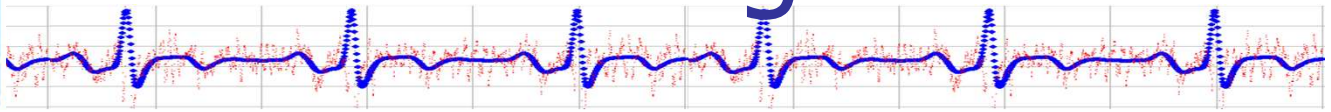
- Question: Is the rate of congressional gun ownership greater than the US national gun ownership rate?
- US: 32% Congress: 55%
  - Based on percentages, yes.
- Are congressional members representative of US households wrt gun ownership?
- If we were to take a random sample of 312 US households, how likely would we be to see this outcome?

Survey Feb 5, 2013 Guns in Congress



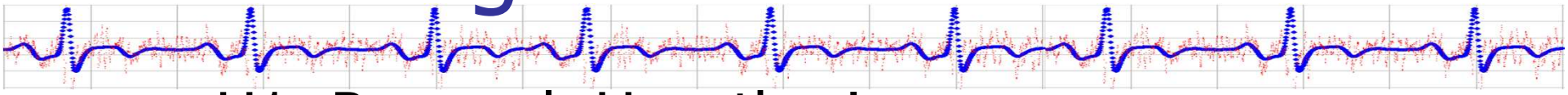
- Null Hypothesis (H_0): distribution of gun ownership in congress = distribution of gun ownership in US households.
- Research Hypothesis (H_1): distribution of gun ownership in congress ≠ distribution of gun ownership in US households.
- Inferential statistics: What is the probability of drawing a sample from US households that has the rate of gun ownership observed in congress?
  - What is $p(H_1|H_0)$?

Survey Feb 5, 2013 Guns in Congress



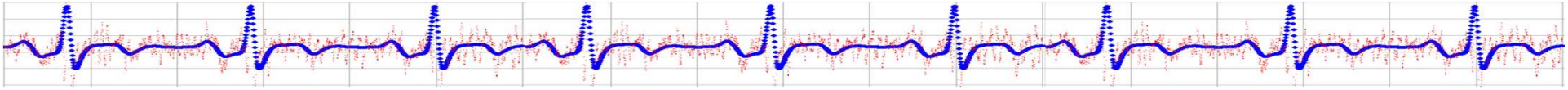
- Significance:
  - How small does $p(H_1|H_0)$ need to be before we say that congress is really different?
  - Convention: $p(H_1|H_0) <= 0.05$
  - "Level of significance" $\alpha = 0.05$
- How to compute $p(H_1|H_0)$?
  - Chi square distribution….

Basic Process of Hypothesis Testing



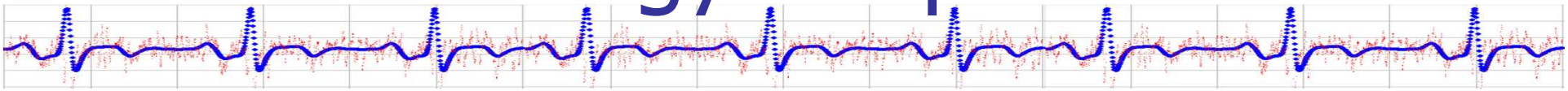
- H1: Research Hypothesis:
  - Population 1 is different than Population 2
- H0: Null Hypothesis:
  - No difference between Pop 1 and Pop 2
  - *The difference is "null"*
- Compute *p(observed difference|H0)*
  - 'p' = probability observed difference is due to random variation
- If p<*threshold* then reject H0 => accept H1
  - p typically set to 0.05 for most work
  - p is called the "level of significance"

Other ways of thinking about this…



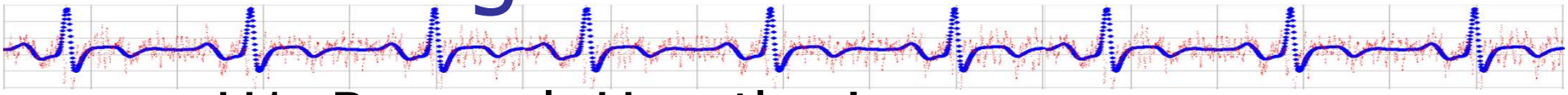
- "Innocent until proven guilty."
- How surprising would this result be if there really were no difference?
- Why do things this way???

Terminology Soup



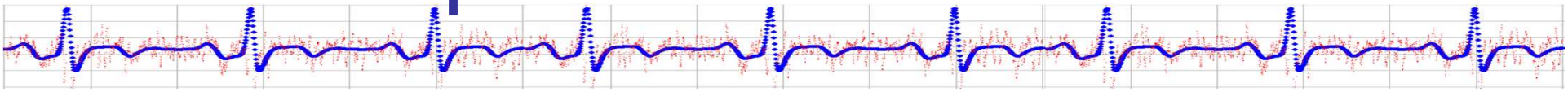
- Descriptive statistics = describing state of the world
- Inferential statistics = Hypothesis testing
  - When you only have access to a sample, or want to account for other kinds of noise
- Univariate analysis/statistics = statistics on only one variable (DV) at a time
- Multivariate analysis/statistics = statistics on more than one variable (DV) at a time

Basic Process of Hypothesis Testing



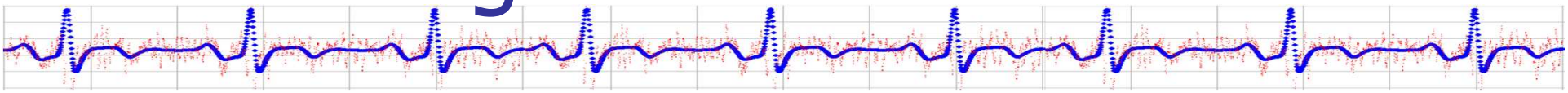
- H1: Research Hypothesis:
  - Population 1 is different than Population 2
- H0: Null Hypothesis:
  - No difference between Pop 1 and Pop 2
  - *The difference is "null"*
- Compute *p(observed difference|H0)*
  - 'p' = probability observed difference is due to random variation
- If p<*threshold* then reject H0 => accept H1
  - p typically set to 0.05 for most work
  - p is called the "level of significance"

Examples



- Research Question:
  - Which is better, Guitar Hero or Rock Band?
- Research Question:
  - Is the ownership of Wii vs. Xbox consoles significantly different among NU students compared to ownership in the general US population?
- Research Question:
  - Are Wii owners more likely to own Guitar Hero vs. Rock Band, compared to Xbox owners?

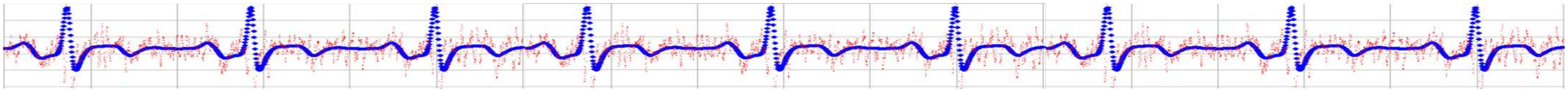
Type of Errors in Hypothesis Testing



Your conclusion	"The Truth" H1 False	"The Truth" H1 True
Accept H1	Type I Error	Correct Decision
Reject H1	Correct Decision	Type II Error

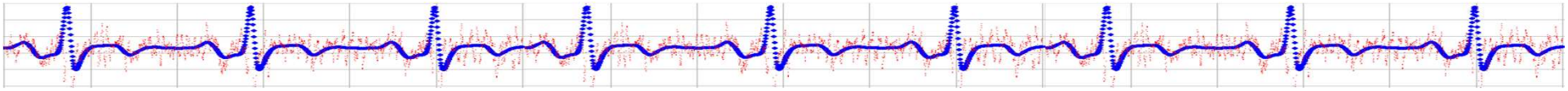
'p' = Probability of Type I Error

Chi-Square for Goodness of Fit



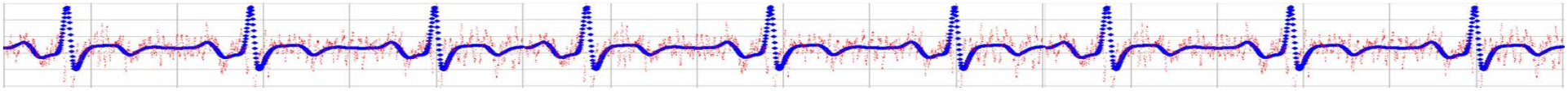
Is an observed frequency distribution significantly different from an expected distribution?

Chi-Square for Goodness of Fit



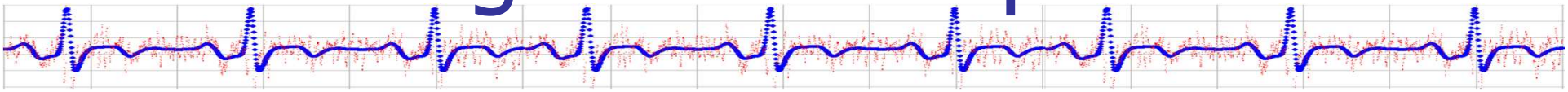
- Assumes
  1. You have a nominal variable
     - Values are exhaustive & mutually-exclusive
  2. You have an *Expected Frequency* table for the nominal variable
  3. None of the expected frequencies are "too small" (≥5)
  4. Random sampling

Chi-Square for Goodness of Fit



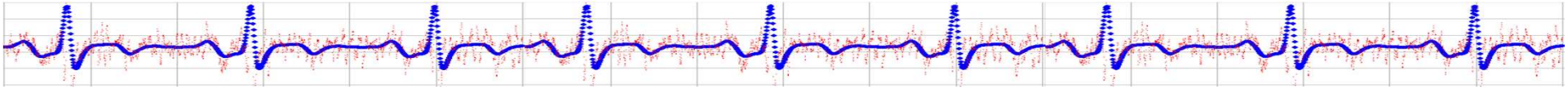
- Form of null hypothesis H0?
  - Observed frequency = Expected frequency
  - Populations (expected, observed) are actually the same on the nominal measure of interest
- Form of hypothesis H1?
  - Observed frequency ≠ Expected frequency
  - Populations (expected, observed) are different

Procedure for Hypothesis Testing with Chi-square



1. Formulate your research hypothesis (including statement of expected frequencies)
2. Determine hypothesis test parameters
   - significance threshold
3. Collect your data
4. Compute Chi-Square statistic and draw conclusion

Formula for Chi-square statistic

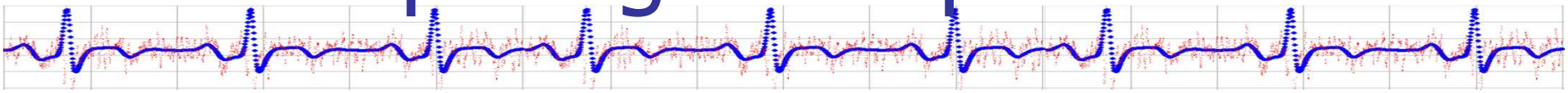


$$X^2 = \sum \frac{(O-E)^2}{E}$$

- O = Observed frequency for a given category
- E = Expected frequency for a given category

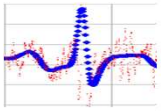
- Note: "statistic" is a function you apply to a set of data (in a statistical analysis)

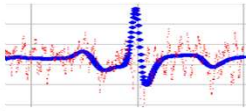
Computing Chi-square

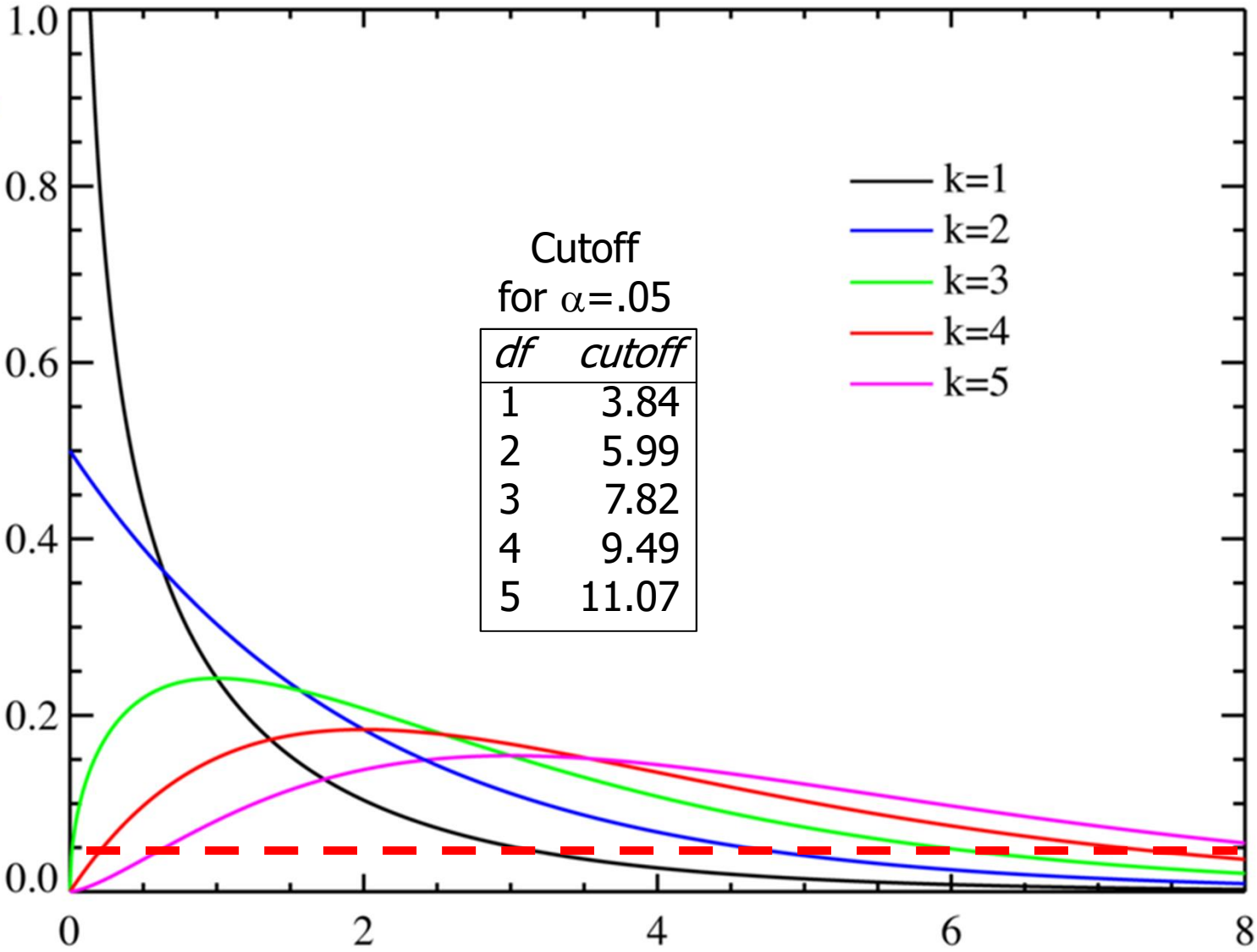


- Manually:
  - Determine df (= num categories – 1)
  - Compute Chi-square using formula
  - Lookup to see if statistic>table entry for significance,df
    - If yes => reject H0
    - If no => inconclusive

Chi-square probability distribution



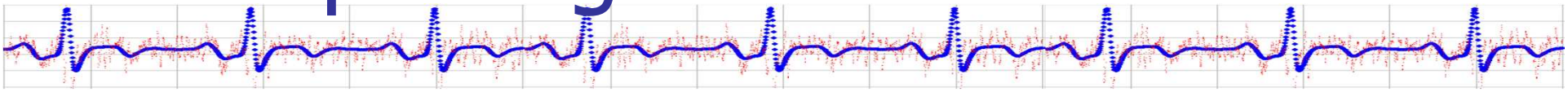




Cutoff for α=.05

df	*cutoff*
1	3.84
2	5.99
3	7.82
4	9.49
5	11.07

Reporting result



$$X^2(df) = chisq, p < sigthresh$$

Where,

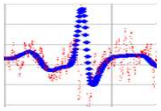
- df = degrees of freedom
- sigthresh = pre-defined significance threshold
  - Note: if p<<sigthresh, can report that as well, e.g., "p<.01", "p=.001"

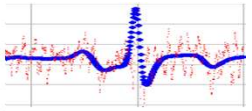
For example: $X^2(2) = 11.89, p < 0.05$

If not significant, than use "n.s." instead of "p<…".

Usually also report expected and actual frequencies, or at a minimum, the total number of cases considered (aka "n").

Chi-square probability distribution





Example

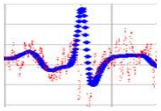
- You go gambling in a shady casino, and suspect that the games are rigged.
- You focus your attention on one 6-sided die being used in a game and keep track of 60 rolls:

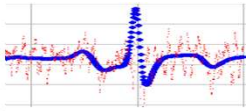
Roll	1	2	3	4	5	6
Count	6	5	7	9	3	30

- Is the die "loaded"?

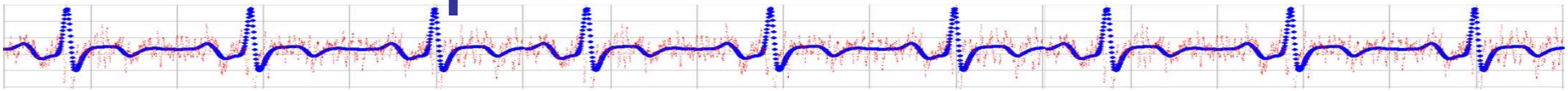
X^2 = 50

Chi-square probability distribution

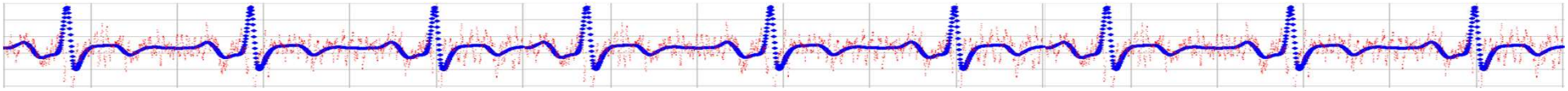




Example



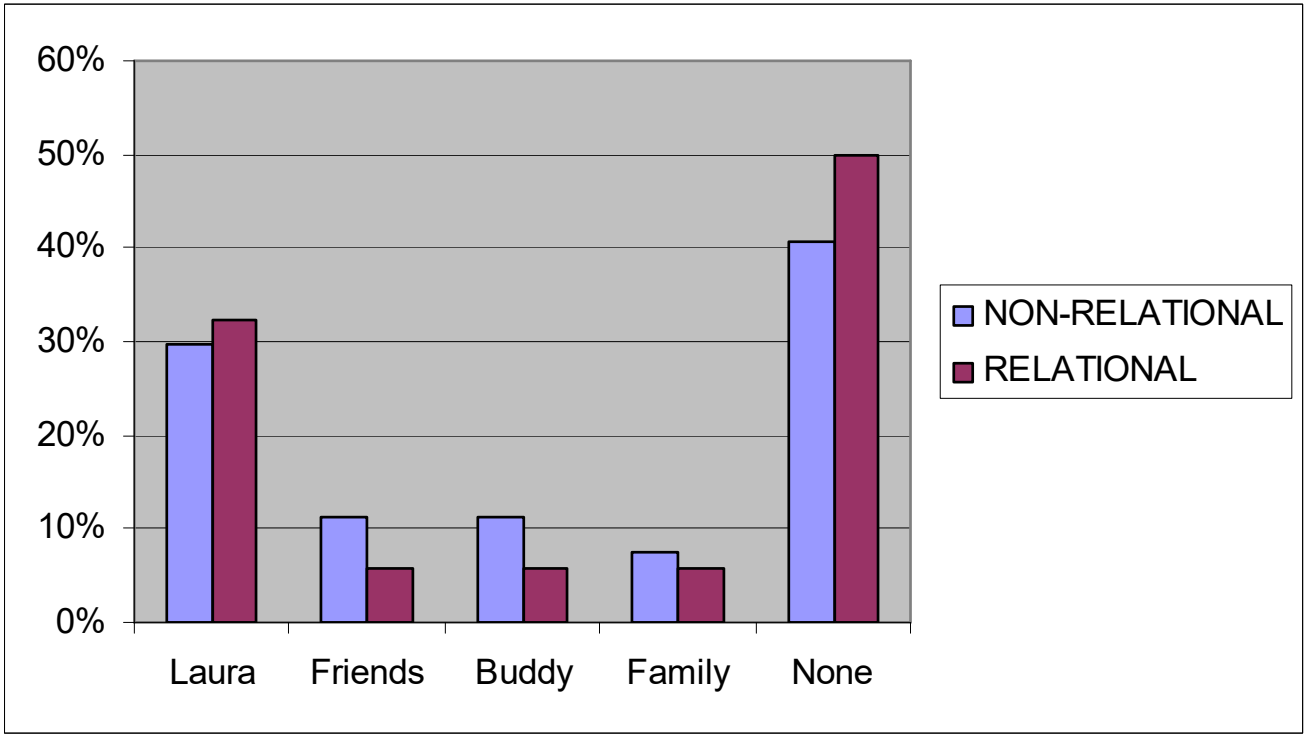
- OS market share in US is:
  - 96% PC; 3% Macintosh; 1% Linux
- You do a survey in your company and find the following user breakdown:
  - 440 PC; 9 Mac; 3 Linux
- Is your company weird?



- X^2 = 2.13, df=2
- P=0.3

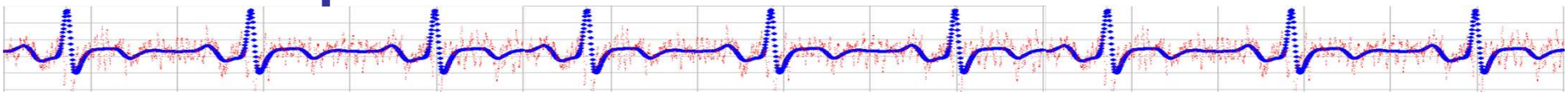
Relational Results
All Subjects

"Who was most helpful in getting you to exercise?"

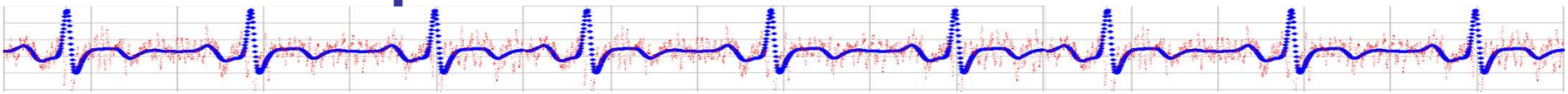


X^2 $p<.05$

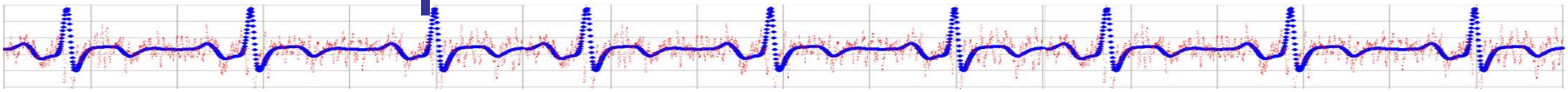
What do you do if you don't have any information to base Expected frequencies on?



Group Exercise



Group Exercise

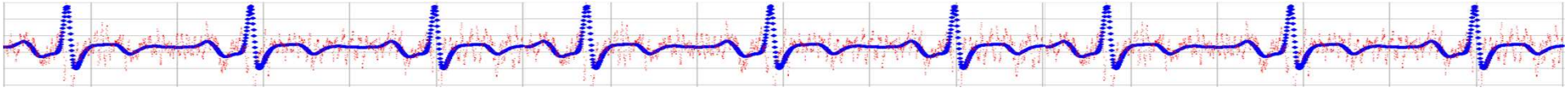


- For each problem, write
  1. Two populations being compared
  2. Research hypothesis
  3. Null hypothesis
  4. Test criteria
  5. Expected frequencies
  6. Observed frequencies
  7. Test results
     - publication format and
     - English

Cutoff for α=.05

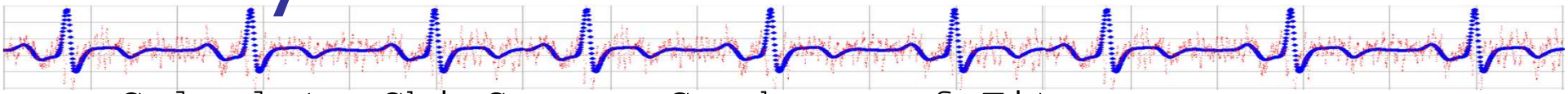
df	*cutoff*
1	3.84
2	5.99
3	7.82
4	9.49
5	11.07

Chi-square Goodness of Fit test



- Which of the following is it appropriate?
  - Descriptive study designs
  - Demonstration study designs
  - Correlational study designs
  - Experimental study designs

Computing Chi-square in Python



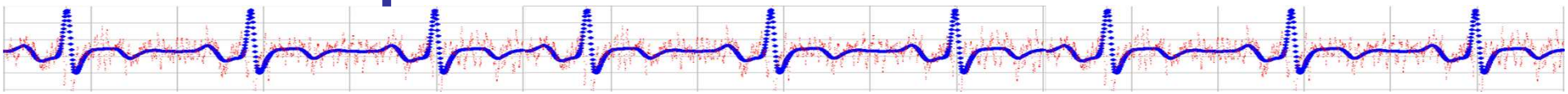
Calculate Chi Square Goodness of Fit

- The results returned are the Chi Square statistic, and the p value. Optionally enter the degrees of freedom, ddof. from scipy.stats
- import chisquare
- observed= [16, 18, 16, 14, 12, 12]
- expected= [16, 16, 14, 15, 13, 12]
- ddof = 2
- chisquare(observed, expected, ddof)

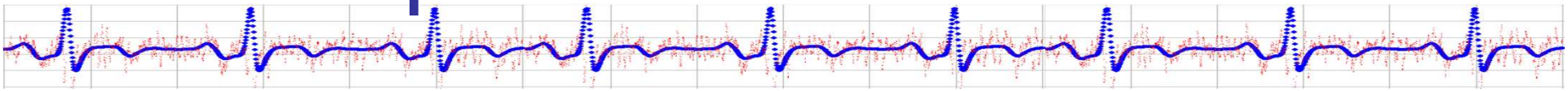
```
>>> (2.0, 0.8491450360846095 6)
```

Note: If you don't supply an expected distribution, it will use a default equal distribution…

Chi-Square Test for Independence



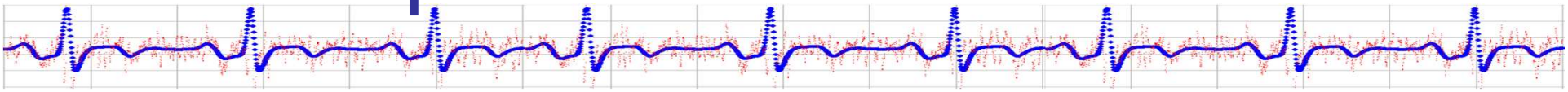
Chi-Square Test for Independence



- Are two variables related (H1), or are they independent (H0)?

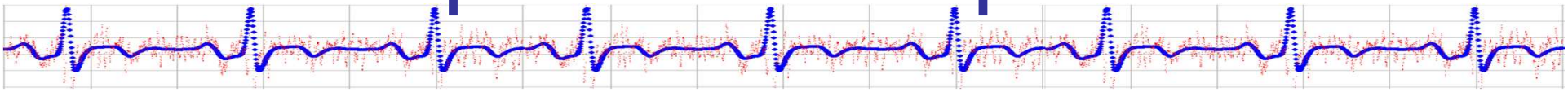
- Assumptons
  - Both variables must be nominal.
  - Cannot be related in a 'special' way (i.e., repeated measures)
  - Random sampling assumed

Chi-square Test for Independence



- Which of the following is it appropriate?
  - Descriptive study designs
  - Demonstration study designs
  - Correlational study designs
  - Experimental study designs

Example from chapter



- Morning & night people using different modes of transportation.
- What kind of table is this?

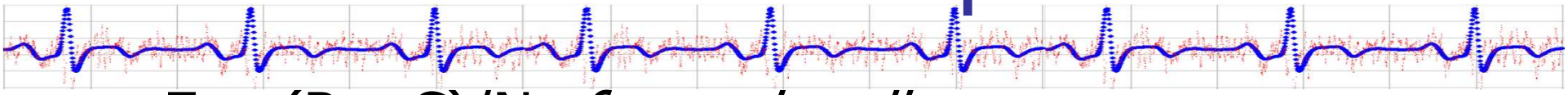
Contingency table

	Bus	Carpool	Own Car
Morning	60	30	30
Night	20	20	40

- What kind of study is this?

Correlational

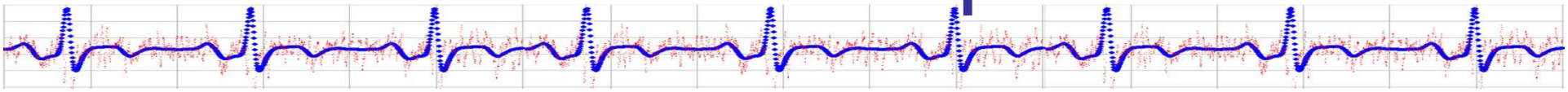
Expected frequencies if variables are independent



- E = (R x C)/N *for each cell*
  - R = row count
  - C = column count
  - N = total number in all cells

	Bus	Carpool	Own Car
Morning	60	30	30
Night	20	20	40

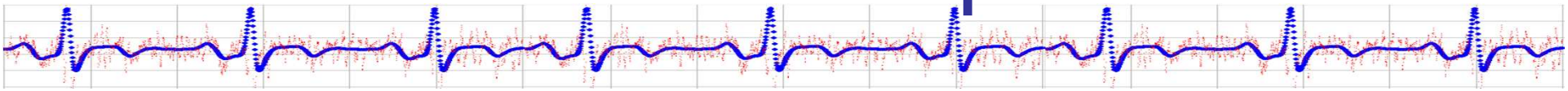
Expected frequencies if variables are independent



- Step 1 – compute row & col totals

	Bus	Carpool	Own Car	
Morning	60	30	30	**120**
Night	20	20	40	**80**
	80	**50**	**70**	

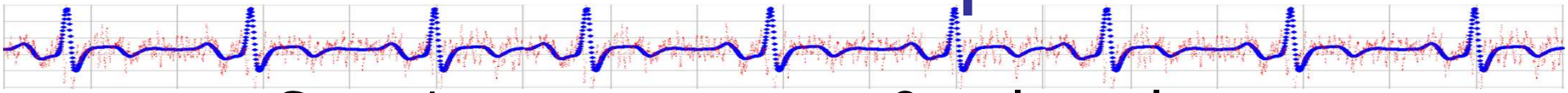
Expected frequencies if variables are independent



- Step 1 – compute row & col totals
- Step 2 – compute row %

	Bus	Carpool	Own Car		
Morning	60	30	30	**120**	**60%**
Night	20	20	40	**80**	**40%**
	80	**50**	**70**		

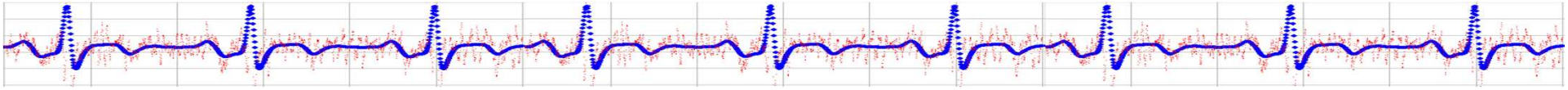
Expected frequencies if variables are independent



- Step 1 – compute row & col totals
- Step 2 – ea cell = (R x C)/N

	Bus	Carpool	Own Car	
Morning	**(48)** 60	**(30)** 30	**(42)** 30	**120**
Night	**(32)** 20	**(20)** 20	**(28)** 40	**80**
	80	**50**	**70**	

Formula

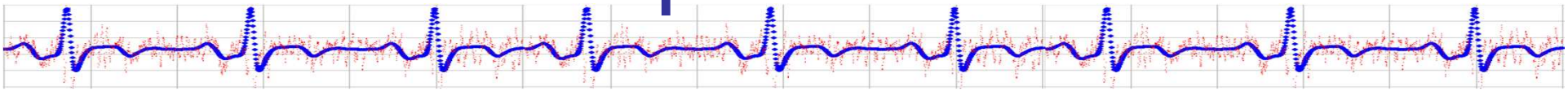


- Same as goodness-of-fit test.

$$X^2 = \sum \frac{(O-E)^2}{E}$$

- df = (NumRows-1)x(NumColumns-1)

Text example



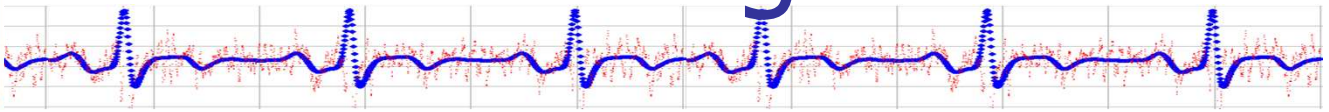
- Chi-Sq = 16.07
- df = ?

Cutoff for α=.05

df	*cutoff*
1	3.84
2	5.99
3	7.82
4	9.49
5	11.07

- Conclusion?

Survey Feb 5, 2013 Guns in Congress

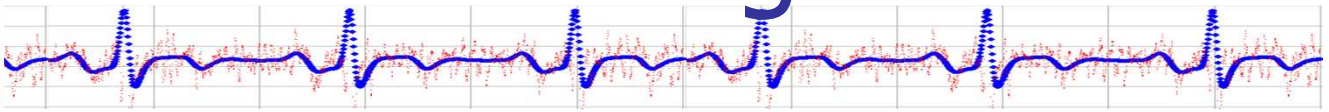


- How many members of congress own guns?
- Dataset:

Title Member Party State GunOwner NRAGrade

- 531 seated members
- 165 Yes, 147 No, 219 “no comment”

Survey Feb 5, 2013 Guns in Congress

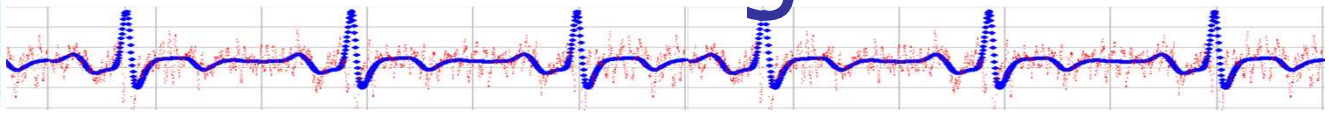


- Given the data, what questions could we ask about the relatedness of nominal measures?

Title Member Party State GunOwner NRAGrade

- Q: Is Gun Ownership related to party?
- Q: Is Gun Ownership related to NRAGrade?

Survey Feb 5, 2013 Guns in Congress



- Q: Is Gun Ownership related to party?

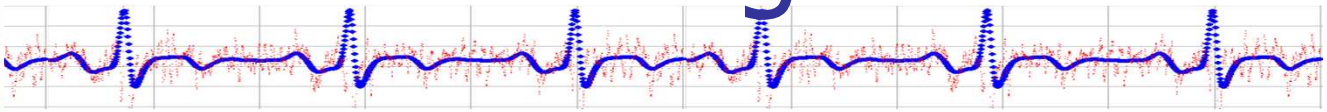
	No	Yes
D	115	44
R	18	119

$df=(2-1) \times (2-1)$

$X^2 = 102$

No Yes A 14 131 B 5 5 C 7 8 D 4 2 F 103 17

Survey Feb 5, 2013 Guns in Congress



- Q: Is Gun Ownership related to NRAGrade?

	No	Yes
A	14	131
B	5	5
C	7	8
D	4	2
F	103	17

df=(2-1) x (5–1)
X^2 = 155