

An automated R tool  
for identifying  
individuals with difficulties  
in a large pool of raters

Pete Meyer and Shaun Lysen  
Google, Santa Monica, California - USA

# Overview

- The User Experience
- How raters assess quality
- Identifying raters that are having difficulties
- Process flowchart
- Summary

# The User Experience

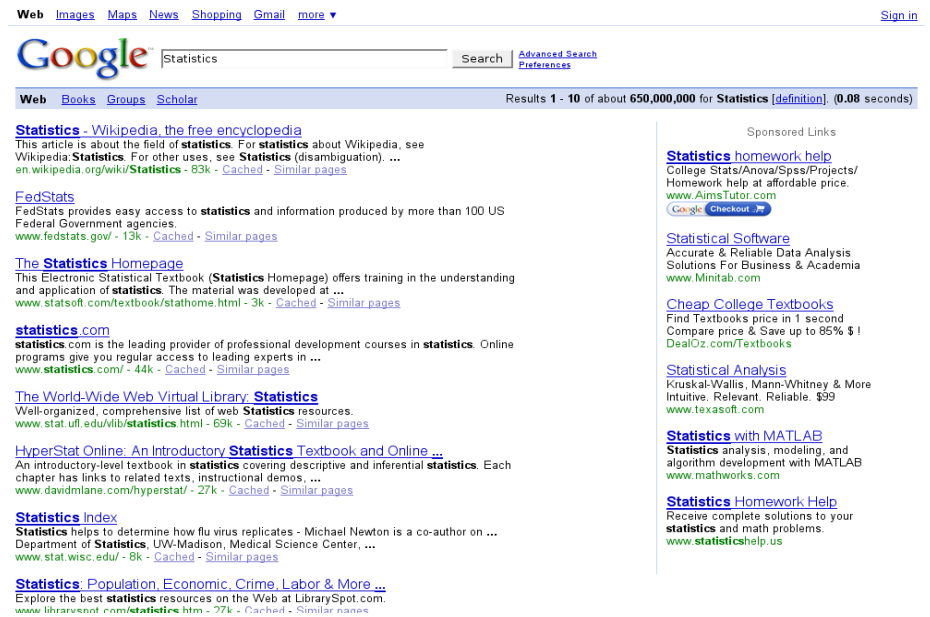
Google's Mission:  
organize the world's  
information and make it  
universally accessible and  
useful.

Google primarily funds the service it provides with  
advertising.

“Eyeballs” drive the value for advertisers.

The User Experience is key to retaining eyeballs.

Ads should contribute to the User Experience, not detract  
from it.



The screenshot shows a Google search for the term "Statistics". The search bar at the top contains the word "Statistics" and a search button. Below the search bar, there are navigation links for "Web", "Books", "Groups", and "Scholar". The search results are displayed in a list format, with the first result being a Wikipedia entry for "Statistics". To the right of the search results, there is a section for "Sponsored Links" which includes several advertisements for statistics-related services and software.

Web Images Maps News Shopping Gmail more ▾ Sign in

Google Statistics Search Advanced Search Preferences

Web Books Groups Scholar Results 1 - 10 of about 650,000,000 for Statistics [definition] (0.08 seconds)

**Statistics** - Wikipedia, the free encyclopedia  
This article is about the field of **statistics**. For **statistics** about Wikipedia, see Wikipedia:**Statistics**. For other uses, see **Statistics** (disambiguation). ...  
[en.wikipedia.org/wiki/Statistics](http://en.wikipedia.org/wiki/Statistics) - 83k - [Cached](#) - [Similar pages](#)

**FedStats**  
FedStats provides easy access to **statistics** and information produced by more than 100 US Federal Government agencies.  
[www.fedstats.gov/](http://www.fedstats.gov/) - 13k - [Cached](#) - [Similar pages](#)

**The Statistics Homepage**  
This Electronic Statistical Textbook (**Statistics** Homepage) offers training in the understanding and application of **statistics**. The material was developed at ...  
[www.statsoft.com/textbook/stathome.html](http://www.statsoft.com/textbook/stathome.html) - 3k - [Cached](#) - [Similar pages](#)

**statistics.com**  
**statistics.com** is the leading provider of professional development courses in **statistics**. Online programs give you regular access to leading experts in ...  
[www.statistics.com/](http://www.statistics.com/) - 44k - [Cached](#) - [Similar pages](#)

**The World-Wide Web Virtual Library: Statistics**  
Well-organized, comprehensive list of web **Statistics** resources.  
[www.stat.ufl.edu/Vlib/statistics.html](http://www.stat.ufl.edu/Vlib/statistics.html) - 69k - [Cached](#) - [Similar pages](#)

**HyperStat Online: An Introductory Statistics Textbook and Online ...**  
An introductory-level textbook in **statistics** covering descriptive and inferential **statistics**. Each chapter has links to related texts, instructional demos, ...  
[www.davidmlane.com/hyperstat/](http://www.davidmlane.com/hyperstat/) - 27k - [Cached](#) - [Similar pages](#)

**Statistics Index**  
**Statistics** helps to determine how flu virus replicates - Michael Newton is a co-author on ... Department of **Statistics**, UW-Madison, Medical Science Center, ...  
[www.stat.wisc.edu/](http://www.stat.wisc.edu/) - 8k - [Cached](#) - [Similar pages](#)

**Statistics: Population, Economic, Crime, Labor & More ...**  
Explore the best **statistics** resources on the Web at LibrarySpot.com.  
[www.libraryspot.com/statistics.htm](http://www.libraryspot.com/statistics.htm) - 77k - [Cached](#) - [Similar pages](#)

Sponsored Links

**Statistics homework help**  
College Stats/Anova/Spss/Projects/  
Homework help at affordable price.  
[www.AimsTutor.com](http://www.AimsTutor.com)  
[Google](#) [Check out](#)

**Statistical Software**  
Accurate & Reliable Data Analysis  
Solutions For Business & Academia  
[www.Minitab.com](http://www.Minitab.com)

**Cheap College Textbooks**  
Find Textbooks price in 1 second  
Compare price & Save up to 85% \$!  
[DealOz.com/Textbooks](http://DealOz.com/Textbooks)

**Statistical Analysis**  
Kruskal-Wallis, Mann-Whitney & More  
Intuitive, Relevant, Reliable. \$99  
[www.texasoft.com](http://www.texasoft.com)

**Statistics with MATLAB**  
**Statistics** analysis, modeling, and  
algorithm development with MATLAB  
[www.mathworks.com](http://www.mathworks.com)

**Statistics Homework Help**  
Receive complete solutions to your  
**statistics** and math problems.  
[www.statisticshelp.us](http://www.statisticshelp.us)



# Raters assess quality

Raters are trained to assign ratings to query-ad pairs according to common guidelines

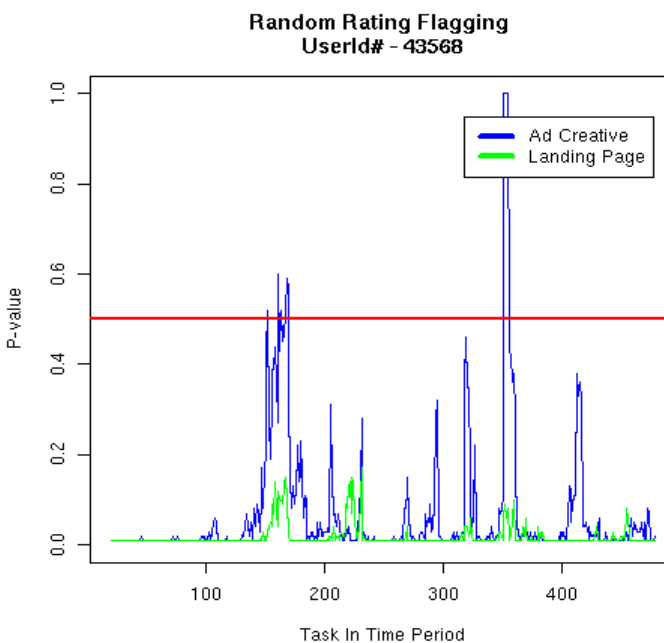
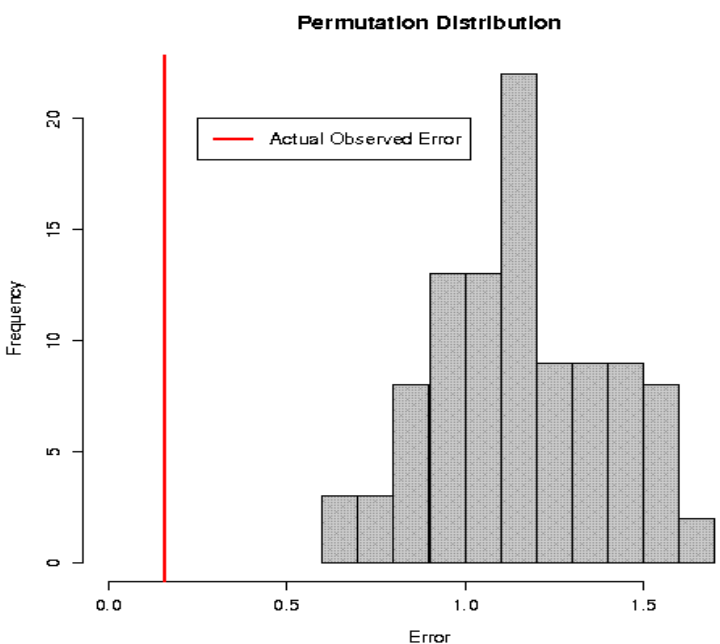
There are a variety of ways raters might diverge from the guidelines, whose detection would require reference to statistical distributions.

- assigning scores randomly
- assigning the same score over and over
- assigning scores without doing due diligence with respect to the landing page
- assigning scores that are inconsistent with the guidelines
- assigning the same score to more than one measure

# Example: Do a series of ratings appear to be random?

## Idea:

Assuming the rater really is rating tasks randomly, then any configuration of his ratings is equally good. Thus under any permutation of his ratings, his error rate should on average be the same.



RatingId	AdCreativeScore	AdCreativeMerit
1848751	1	1.75
1848753	1	1.56
1848754	1	1.44
1848756	1	-0.75
1848757	1	0.50
1848758	1	2.12
1848759	1	1.06
1848760	1	1.31
1848761	1	1.50
1848762	1	1.31
1848763	1	1.38
1848765	1	0.44
1848766	1	1.88
1848767	1	0.69
1848769	1	1.50
1848771	1	0.50
1848773	1	1.56
1848774	1	0.38
1848775	1	0.94
1848776	1	1.2

## Example:

Are unusually long runs of the same score assigned?

## Idea:

Given the proportions of each rating occurring over a week and the number of ratings submitted for a given rater, how unusual is it to see run lengths as long as those observed?

## Simulated run lengths:

<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>	<u>11</u>
353289	52483	9511	1914	437	87	21	5	2	0	1

## Longer observed run lengths:

<u>12</u>	<u>13</u>	<u>14</u>	<u>15</u>	<u>18</u>	<u>22</u>	<u>25</u>
6	1	1	1	2	1	1

# Notifying managers

## Construct an HTML results file and send a plain text email

Rater flags for ratings from 2008-07-13 to 2008-07-21

Report generated Mon Jul 21 08:24:03 2008

Rater flags - ordered by UserId within Preferred Language

UserId	Login	Language	FlagId	Comment
21217			1	Run of 38 (PageToAdScoreFirst=Excellent)
25246			1	Run of 14 (LandingPageScoreFirst=DissatisfactionLikely)
26603			1	Run of 24 (PageToAdScoreFirst=Excellent)
26603			1	Run of 32 (PageToAdScoreFirst=Excellent)
32297			1	Run of 11 (AdCreativeScoreFirst=SatisfactionPossible)

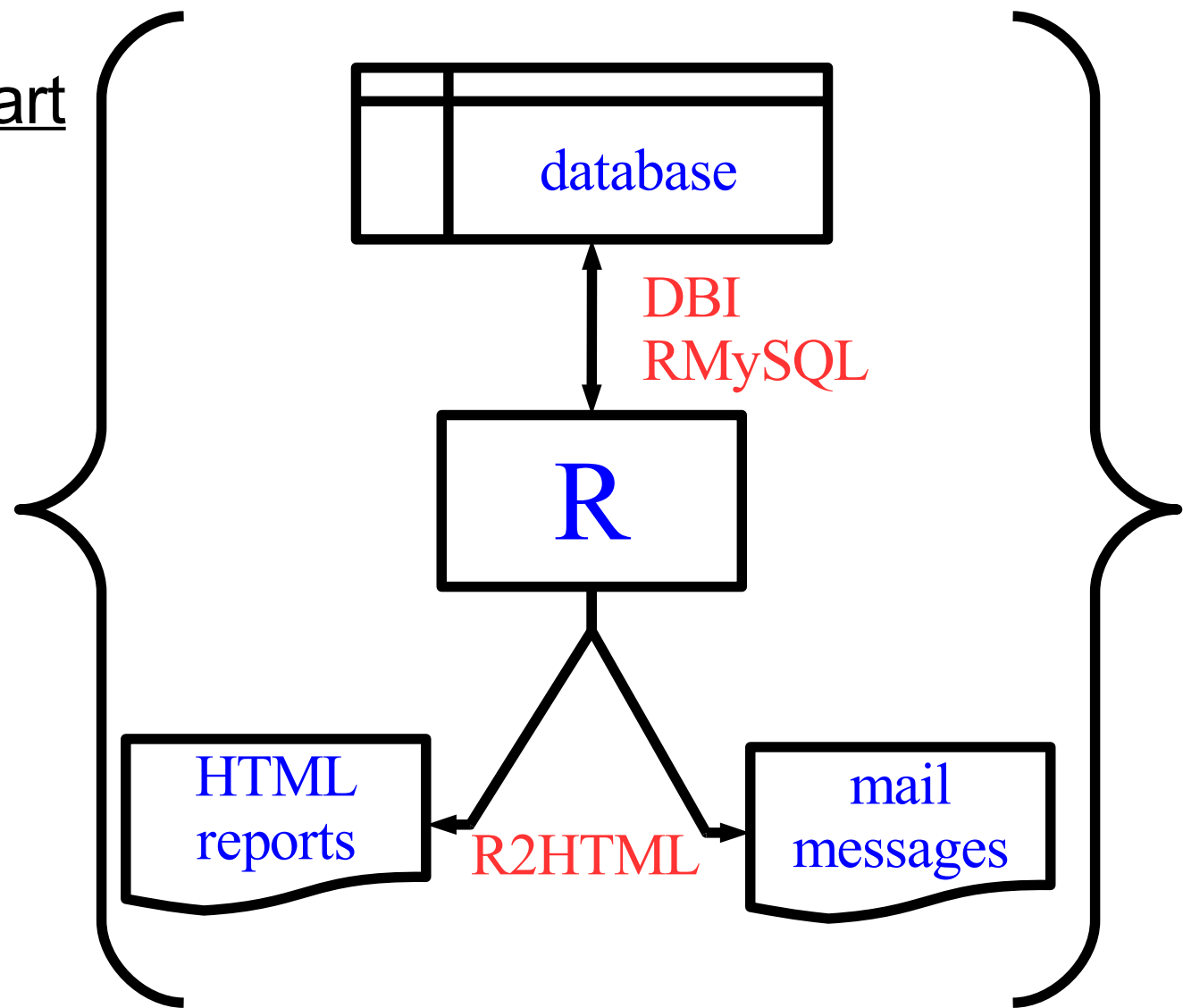
```
system(paste('mail -s',subj, ' ',paste(recipients, collapse=','),' < temp0001.txt', sep=''))
```

## Send an HTML email

```
paste("mutt -e 'set content_type=\"text/html\"'", paste(recipients, collapse=","),  
      "-s", paste("","subj, """, sep=""), "<", fileName)
```

# Process flowchart

crontab



```
# m h dom mon dow  command
1 2 * * 1 . <home directory>/bashrc; R --vanilla < RaterFlagging-6.R
```



# Credits (and many thanks!) go to ...

R Core

DBI: R-Databases Special Interest Group

RMySQL: David A. James <dj@bell-labs.com>  
Saikat DebRoy <saikat@stat.wisc.edu>

R2HTML: Eric Lecoutre



# Summary

R (with DBI, RMySQL, and R2HTML)

enabled us to leverage statistical insights

that are not accessible through standard database  
tools

in order to identify raters that are having difficulties

and communicate the results to colleagues

in a production environment.