

Matt Adelman
Dan Hemmendinger
Brian Stellingwerf

The Wikipedia Game Experiment

Is there a difference in the mean amount of clicks between males and females in the Wikipedia Game?

In order to determine the origin and to compute the data for this experiment the question must first be examined. The mean amount of clicks presented in the experiment is the total clicks on the computer it takes to get from the starting point on a given Wikipedia page to the ending point (another given site). The directions given to the subjects were as follows:

The objective of this game is simple. Go from one starting Wikipedia article to the given ending article, using only the hyperlinks within the articles (no using the links on the left bar) as you go. You are not allowed to click “Backspace” or go forward and back in articles.

Remember, the object of the game is to get from the link provided to the stated article in the fewest number of *clicks*, it is important to count the number of clicks you use. Also, a click that takes you to a location within the same article does not count as a click. Curl (the find feature) is also prohibited.

Please use the provided score sheet to mark yourself.

Scroll down for a practice round.

Each individual round includes these instructions to all of the subjects:

Round (1, 2, 3, or 5) {Round 4 was skipped due to lack of time}

From: (Some given website)

To: (Some given word)

When finished, please wait for instructions.

DO NOT SCROLL DOWN!!

After the round was completed, all of the subjects moved on together. The four rounds of starting sites were given to the subjects individually and they were as follows:

Round 1

From:

<http://en.wikipedia.org/wiki/Incognito> (Amanda Lear album)

To:

When finished, please wait for instructions.

DO NOT SCROLL DOWN!!

Round 2

From:

http://en.wikipedia.org/wiki/Ducs_d%27Angers

To:

When finished, please wait for instructions.

DO NOT SCROLL DOWN!!

Round 3

From:

http://en.wikipedia.org/wiki/Nea_Filadelfeia

To:

When finished, please wait for instructions.

DO NOT SCROLL DOWN!!

Round 5

From:

http://en.wikipedia.org/wiki/Guildford_railway_station,_Sydney

To:

When finished, please wait for instructions.

DO NOT SCROLL DOWN!!

The section that says “To:” was not filled in so that the subjects would not skip ahead, since all of the sites were given to the subjects at once but they were completed one at a time. When the round began and the link was clicked, the subjects wrote the corresponding words into their sheets in the “To:” spots. We, as researchers, had the “cheat sheet” of answers that was kept between the three of us which contained the “To:” answers as follows:

Round 1

To: Maple Syrup

Round 2

To: Water

Round 3

To: Ulysses S. Grant

Round 5

To: Cactus

The subjects were told to write down every step in order to be able to track the exact results and in order to prevent any cheating or more confounding variables. The score sheets looked similar to the one below along with two examples, one original with no writing, one from a male, and one from a female.

Wikipedia Game Experiment Score Sheet

Round
Practice To <u>Dark Matter</u>
1 To _____
2 To _____
3 To _____
4 To _____
5 To _____

Write the steps you took, and total number of clicks for each round to the right of the round number.

Age: 16
Gender: Male

Wikipedia Game Experiment Score Sheet

Round
Practice To <u>Dark Matter</u>
1 To <u>Maple Syrup</u>
2 To <u>Water</u>
3 To <u>Ulysses S Grant</u>
4 To _____
5 To <u>Cactus</u>

Write the steps you took, and total number of clicks for each round to the right of the round number.

Antarctic Ocean ¹ → Antarctica ² → Earth ³ → mass ⁴ → matter ⁵ → Dark Matter
5 clicks

Amardis ¹ → France ² → French Language ³ → Canada ⁴ → Flag of Canada ⁵ → Maple Leaf
→ Maple → Maple Syrup
7 clicks

Ice hockey → Ice ¹ → Water (Molecule) ² → Water ³
3 clicks

Philadelphia ¹ → Philadelphia (Penn.) ² → Philadelphia, 1776 ³ → United States ⁴
→ American Civil War ⁵ → Ulysses S. Grant ⁶
6 clicks

X

→ South line, Sydney ¹ → Australia ² → Australia (continent) ³ → English ⁴ → English Language
→ United States ⁵ → Colorado ⁶ → Cactus ⁷
7 clicks

Female 17

Wikipedia Game Experiment Score Sheet

Write the steps you took, and total number of clicks for each round to the right of the round number.

Round
Practice To <u>Dark Matter</u>
1 (8) To <u>Maple Syrup</u>
2 (3) To <u>Water</u>
3 (6) To <u>Ulysses S. Grant</u>
4 To _____
5 (14) To <u>Cactus</u>

French France North America Canada Flag of Canada maple leaf maple tree
maple syrup

ice hockey ice water

Greece Greek civil war World War II United States American Civil War
Ulysses S. Grant

south railway line, sydney, Australia, desert, Great Sandy desert, spinifex
austrailian, outback, the bush, Canadian French, United States
Great Plains, New Mexico, cacti

X

Each of the words for the end of the search was also randomly chosen by the researchers. Once the starting and ending points were chosen at random, the chain between each of the starting and ending points were checked to guarantee a chain between them. This assures that the experiment is possible and valid. Then, the experiment commenced and the results were calculated using a two-sample *t*-test.

Two-Sample *t*-Test

Original Data:

Males:

Subject	Average Clicks
Male 1	5.75
Male 2	6
Male 3	5.75
Male 4	5.75
Male 5	4.666667
Male 6	4.25
Male 7	3.75
Male 8	3.5
Male 9	3.333333
Male 10	4
Male 11	5.5

Females:

Subject	Average Clicks
Female 1	5.25
Female 2	6.75
Female 3	5.75
Female 4	9
Female 5	6
Female 6	5.5
Female 7	7.75
Female 8	5.5
Female 9	6.25

Range: 2.666667

Min: 3.333333

Max: 6

Q1: 3.75

Q3: 5.75

Median: 4.666667

Range: 3.75

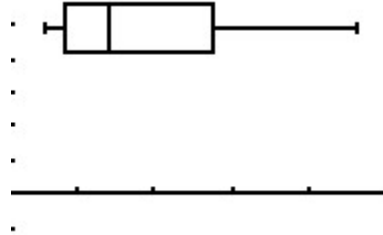
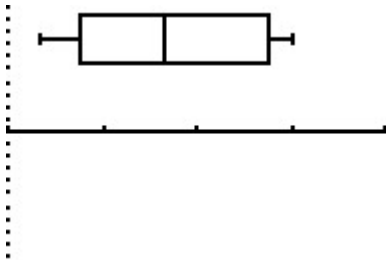
Min: 5.25

Max: 9

Q1: 5.5

Q3: 7.25

Median: 6



μ_1 = The population mean amount of clicks for males.

μ_2 = The population mean amount of clicks for females.

\bar{x}_1 = The sample mean amount of clicks for males.

\bar{x}_2 = The sample mean amount of clicks for females.

S_1 = The sample standard deviation for males.

S_2 = The sample standard deviation for females.

n_1 = The sample size for males.

n_2 = The sample size for females.

t = The test statistic.

t* = The critical value at which the null hypothesis will be rejected if surpassed by the test statistic.

α = The level of significance.

df = The degrees of freedom (found using technology).

p-value = The probability that the next sample is as extreme or greater than the current sample. It is doubled because of the nature of the “Two-Tailed” test (found using the *t* cumulative density function).

H₀: Males and females have the same mean amount of clicks ($\mu_1 = \mu_2$).

H_a: Males and females have different mean amount of clicks ($\mu_1 \neq \mu_2$).

α = 0.05

df = 15.4775

|t* = 2.13145 (found using 15 degrees of freedom)

\bar{x}_1 = 4.75

\bar{x}_2 = 6.4167

S₁ = 1.0254

S₂ = 1.2374

n₁ = 11

n₂ = 9

Assumptions:

Groups are independent

Data in each group are independent

Both populations are approximately Normal

The sample size constitutes less than 10% of the total population

Simple Random Sample (this assumption is violated which means the results of this test need to be interpreted with caution)

Test Statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)}}$$

Decision Rule:

If $|t^*| > |t|$ then the null hypothesis has failed to be rejected, however if $|t^*| < |t|$ then the null hypothesis is rejected.

Calculations:

$$t = \frac{(4.75 - 6.4167) - (0)}{\sqrt{\left(\frac{1.0254^2}{11} + \frac{1.2374^2}{9}\right)}} \qquad t = \frac{\left(-\frac{5}{3}\right) - (0)}{\sqrt{\left(\frac{1.0514^2}{11} + \frac{1.5315^2}{9}\right)}}$$

$$t = \frac{\left(-\frac{5}{3}\right)}{\sqrt{(0.9558 + 0.1701)}} \qquad t = \frac{\left(-\frac{5}{3}\right)}{\sqrt{(0.2657)}} \qquad t = \frac{\left(-\frac{5}{3}\right)}{(0.5155)}$$

$$t = -3.2332$$

$$p - value = 2 \times t - cdf(-\infty, -3.2332, 15) \qquad p - value = 0.0056$$

$t = |-3.3781| > |2.13145| = t^*$ therefore the null hypothesis is rejected.

At the $\alpha = 0.05$ level of significance, there is sufficient evidence to suggest that there is a significant difference in the mean amount of clicks in the Wikipedia Game between males and females.

Two-Sample Confidence Interval

To confirm the previous results, a 95% confidence interval was performed.
Assumptions from Two-Sample t -test

Assumptions:

Groups are independent

Data in each group are independent

Both populations are approximately Normal

The sample size constitutes less than 10% of the total population

Simple Random Sample (this assumption is violated which means the results of this test need to be interpreted with caution)

Formula:

$$(\bar{x}_1 - \bar{x}_2) \pm t * \left(\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \right)$$

Calculation of Interval:

$$(4.75 - 6.1467) \pm 2.13145 * \left(\sqrt{\frac{1.0254^2}{11} + \frac{1.2374^2}{9}} \right)$$

$$\left(-\frac{5}{3} \right) \pm 2.13145 * \left(\sqrt{\frac{1.0514}{11} + \frac{1.5313}{9}} \right)$$

$$\left(-\frac{5}{3} \right) \pm 2.13145 * (\sqrt{0.9558 + 0.1701})$$

$$\left(-\frac{5}{3} \right) \pm 2.13145 * (\sqrt{0.2657}) \quad \left(-\frac{5}{3} \right) \pm 2.13145 * (0.5155)$$

$$\left(-\frac{5}{3} \right) \pm 1.0988 \quad -2.765 \leq (\mu_1 - \mu_2) \leq -0.568$$

This means it can be said with 95% confidence that there **is** a significant difference in the mean number of clicks between males and females in the Wikipedia Game because the interval does not include the value of 0. Since this interval is on the negative side, it would also be appropriate to say that males recorded less clicks than the females because it was the males that were \bar{x}_1 and the smaller the mean, thusly, fewer clicks were used by the males than females.

After analyzing all of the results given from the *t*-test and the confidence interval, the confounding variables and results were examined to make future investigations and experiments on this subject easier and more efficient. There were several confounding variables that contributed to the Wikipedia Game Experiment which could have contributed to different data, or the variables might have had no effect at all.

Confounding Variables

One major confounding variable was the observed subversion by one particular student. Some of the actions taken by this student include the copying of others, not participating, and attempting to complete other tasks. Cheating was also observed with a few other subjects, in the form of answer sharing and reattempting the game once complete. These confounding variables made the answers less trustworthy; however, these disruptions likely did not affect the entire sample as cheating was common among both males and females.

Strengths/Changes

There were several strengths which allowed for the experiment to get decent results such as the unknown ending point, which allowed the subjects not only to concentrate on the given round but also to be more efficient and work alone rather than in groups. The changes came mostly from the confounding variables. Among all of the negatives and confounding variables the experiment presented, there were innumerable amounts of positive outcomes which were implemented as well. What would be changed from the original experiment should be:

- Larger sample size by asking more people to participate with a true Simple Random Sample of all seniors
- Isolation of each subject to avoid some of the confounding variables
- Silence
- Knowledge of the smallest route between the start and end point should not be known among subjects
- Make sure that everyone has the required materials (i.e. – pencils) before the experiment begins in order to save time
- Prevent the talking and hounding of subjects on other subjects (refers to the silence and isolation of each subject)
- No recordings should take place among spectators of the experiment
- Preparation time should be maximized before the subjects enter as to not waste time or add a variable to the experiment
- Motivation should be used in order to get full attention and focus on the experiment

New Experiment/Further Research

Another selection of research that is similar to the method used is the usage of time rather than number of clicks to determine the distribution. By using time, this would prevent the subjects from going back to the starting point and trying to find the shortest route, as the route will not matter, as long as they get the test completed in the fastest time. While this research is more difficult to track for a large amount of subjects, if they were each tested individually, or even in smaller groups, it would be possible to time, track, and analyze the different times into a similar *t*-distribution.