

wikimetrics
metrics.wmflabs.org

Revised May 2015



Agenda

1. Overview
2. How Wikimetrics works
3. Metrics
4. Gathering usernames
5. Demo!
6. Demo!



What is Wikimetrics?

- Online API for measuring user contributions

What it offers:

- Standardized measures of data (“metrics”)
- Open and easy access to data
- Cohort analysis



Why Wikimetrics?

User Characteristics. How many users at this event are new? How many are existing active editors (made 5 edits last month)?

User Activity. How many articles and how many bytes did a group of users produce during an event?

User Retention. Are they still editing 1, 3, 6 months later? Who is still editing?



Some key terms

Cohorts

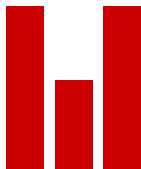
A group of users sharing one or more property or attribute, such as time of account creation or participating in the same program implementation.

example: Users who participated in an edit-a-thon in July.

Metrics

Well-defined values or sets of values that can be computed for any registered user

example: The number of times users clicked the “save page” button



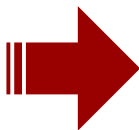
How Wikimetrics Works



Cohorts

My list of Usernames

Iskdjfoid
laksidghesop
oaitoinsldkn
asdnvoasdn



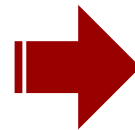
Wikimetrics

Welcome to the Wikimedia Foundation's Wikimetrics homepage. This API allows you to select a set of users, also known as a 'cohort' (for example, all users who signed up via the Thank You campaign)

select a metric to be computed for each of these users (for example, how many [bytes they've added](#)) with optional parameters (for example, a time range) and retrieve the response in JSON or CSV format.

You can also compute a single, aggregate value for the cohort (like the mean revert rate) .

[Learn More](#) [Analyze](#)



Metrics

username	bytes added
Iskdjfoid	654
laksidghesop	0
oaitoinsldkn	456
asdnvoasdn	7895



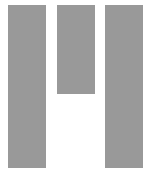
Log in!

Log In

To use Wikimetrics, you must log in with one of the following services:



What does it mean to log in with a service above? The method we use is called [OAuth](#). In plain words, it allows a provider to vouch that you are who you say you are. For example, if you click on "log in with Google" above, Google will ask you if you will allow the Wikimetrics application access to your private Google data. For our purpose, we will only access your email address and name to verify your identity. We will not share any of your data with Google or anyone else. Since Wikimetrics is an open source project, you can see exactly how we interact with Google and the other providers. As of this writing, that logic [lives here](#).



Upload/Create a Cohort

Cohort
My list of Usernames
 Wikimaniaisfun, enwiki
 YayLondon, eswiki
 BarbicanPower, dewiki
 WP4LYFE, ptwiki

Create a Cohort

Name
(automatically checks availability)

Description

Default Project

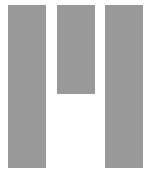
Upload File No file chosen

OR

Paste Usernames

CentralAuth Expand cohort to all user's accounts across projects (via CentralAuth)

Validate As User Ids (Numbers found in the user_id column of the user table)
 User Names (Names found in the user_name column of the user table)



Check for validation

PRGRM_ENTITY_EVENT_DATE

PGRM2_ENTITY2_EVENT2_DATE2

asdfasdf

EDITHN_WMSHSTRY_RSEdin_3DEC13asdfasdf

asdfasdfasdfewrt2t345yqgadfgzxf

TESTTESTTEST

NYCCOHORTSAMPLE

PRGRM_ENTITY_EVENT_DATE A

description of the cohort

sample-tag x you-can-delete-them! x

Add Tag

Validated 116 of 116 cohort members. 71 are valid, 45 are invalid.

Create Report

Remove Cohort





Create a Report

Create New Report



Create Report

Make this a **Public Scheduled** Report. This means that it will run daily at 00:00 UTC and compute results for each day it runs.

Checking this box makes the results of this report **publicly** accessible. There will be no way to stop the report once it starts. When using this feature, make sure you understand these caveats or contact wikimetrics@lists.wikimedia.org for help.

Pick Cohorts



- testestsets
- PRGM_ENTITY_EVENT_DATE

Pick Metrics

- Bytes Added
- Edits
- Newly Registered
- Pages Created

Compute different aggregations of the bytes contributed or removed from a mediawiki project

Pick your Metrics



Pick Metrics

- Bytes Added
- Edits**
- Newly Registered
- Pages Created
- Rolling Active Editor
- Survival
- Threshold

Compute the number of edits in a specific namespace of a mediawiki project

Start Date: 2014-06-29 00:00:00

End Date: 2014-07-29 00:00:00

Time Series by: none

Namespaces: 0 (comma separated)

Refers to the division of pages by namespace in wiki projects. More information:

- [What is a namespace?](#)
- [Universal namespace codes](#)
- Examples: [dewiki](#), [commons](#), [enwiki](#)

Validate Configuration

add parameters!

Configure Output

“Configure output” means how you want your data to be presented in the report.

You can choose **Individual results**, or data per user, or Group level (i.e. **Aggregate results**), which sums up all users, or both!

Configure Output

Bytes Added

Individual Results

Aggregate Results

Sum

Average

Standard Deviation



Sample Result

Bytes Added - testtestsets

Name

Project

Bytes Added

Sum



Run Report!

Run Report

Current Report Inbox

This table refreshes automatically every 5 seconds. Reports **that are not recurrent** will expire after 30 days, please save the output if you would like to keep it longer.

Create New Report



Bytes Added - testestsets

Today at 00:18 UTC

SUCCESS

Get Data ▾

Entypo icons by [Daniel Bruce](#)

CSV and JSON

Comma-separated value file

```
user_id,user_name,negative_only_sum,net_sum,positi
Sum,,-23.0000,-23.0000,0.0000,23.0000
,,,,,
,,,,,
parameters,,,,,
Cohort,TESTTESTTEST,,,,
Cohort Size,3,,,,
Created On,2014-05-27 14:52:01,,,,
Metric,BytesAdded,,,,
Metric_absolute_sum,True,,,,
Metric_aggregateAverage,False,,,,
Metric_aggregateResults,True,,,,
Metric_aggregateStandardDeviation,False,,,,
Metric_aggregateSum,True,,,,
Metric_end_date,2014-05-27 00:00:00,,,,
Metric_id,bytes-added,,,,
Metric_individualResults,False,,,,
Metric_namespaces,0,,,,
Metric_negative_only_sum,True,,,,
Metric_net_sum,True,,,,
Metric_outputConfigured,True,,,,
Metric_positive_only_sum,True,,,,
Metric_start_date,2014-04-27 00:00:00,,,,
Metric_timeseries,none,,,,
```

JavaScript Object Notation

```
"result": {
  "Sum": {
    "negative_only_sum": -23.0,
    "net_sum": -23.0,
    "positive_only_sum": 0.0,
    "absolute_sum": 23.0
  }
},
"parameters": {
  "Cohort": "TESTTESTTEST",
  "Metric_absolute_sum": true,
  "Metric_positive_only_sum": true,
  "Metric_id": "bytes-added",
  "Metric_timeseries": "none",
  "Metric_aggregateResults": true,
  "Created On": "2014-05-27 14:52:01",
  "Metric": "BytesAdded",
  "Metric_net_sum": true,
  "Metric_individualResults": false,
  "Metric_start_date": "2014-04-27 00:00:00",
  "Metric_negative_only_sum": true,
  "Metric_namespaces": "0",
  "Metric_aggregateStandardDeviation": false,
  "Metric_aggregateSum": true,
  "Cohort Size": 3,
  "Metric_end_date": "2014-05-27 00:00:00",
  "Metric_aggregateAverage": false,
  "Metric_outputConfigured": true
}
```



How to collect usernames

Online

- Event pages on-wiki (i.e. contests, edit-a-thons)
- Course pages (Outreach extension)
- Registration

Offline

- Sign-up sheets from in-person events



Consent to use Wikimetrics

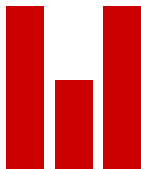
Online

- Include an "opt-in" or "opt-out" clause to allow participants to excuse themselves from being included in wikimetrics cohort

Offline

- Must ask for "opt-in" consent before including username in Wikimetrics cohorts

For correct language, see: [Opt-in forms](#)



Metrics

- **Bytes**
- **Edits**
- **Newly Registered**
- **Pages Created**
- **Rolling Active Editor**
- Rolling New Active Editor
- Rolling Surviving New Active Editor
- Threshold
- Survival



(demo)

Bytes added

The amount of content a user/cohort has added or removed within a given time period.

Net sum Bytes added minus bytes removed

Absolute sum Number of bytes added + number of bytes removed

Positive only sum All bytes added

Negative only sum All bytes removed



(demo)

Edits

The number of times a user/cohort clicked the “save page” button

By clicking the "Save page" button, you agree to the [Terms of Service](#) and have the understanding that a hyperlink or URL is sufficient for CC BY-NC-SA.

Save page

Show preview

Show changes

Cancel

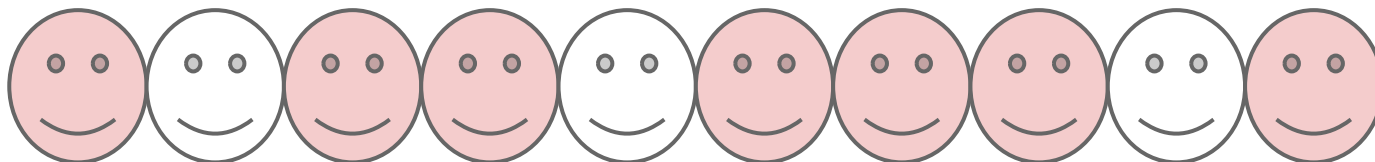
Doesn't include number of bytes per edit!



(demo)

Newly Registered

Which users in the cohort are newly registered between the start and end time





(demo)

Pages Created

The number of new pages
created by a user/cohort

In Commons, uploading a new media file
(image, movie, audio) creates a page!



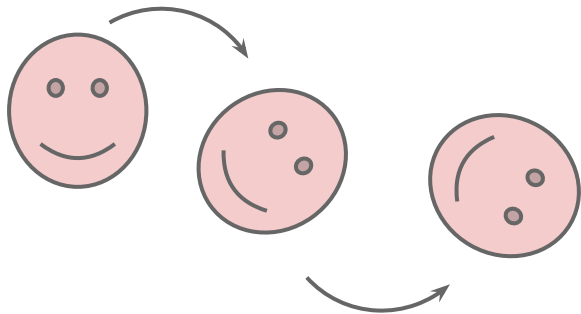
Rolling Active Editor

Number of **registered** users with n edits on any page of a WM project between an end date t and d number of days *before*.

30 days prior
(d)

May 14, 2014
(t)

Did they make 5 edits during this time?





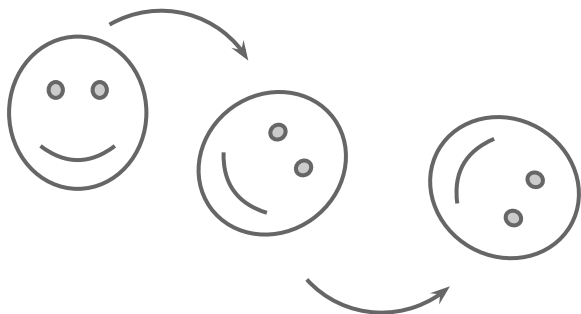
Rolling New Active Editor

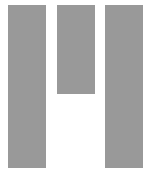
Number of **new user accounts** in a cohort with n edits on any page of a WM project between an end date t and d days *before*.

30 days prior
(d)

August 1, 2014
(t)

Did they make 5 edits during this time?





Rolling Surviving New Active Editor

Number of **new user accounts** with n edits on any page of a WM project during both:

- between an end date t and d days *before*
- between an end date t and $2*d$ days *before*

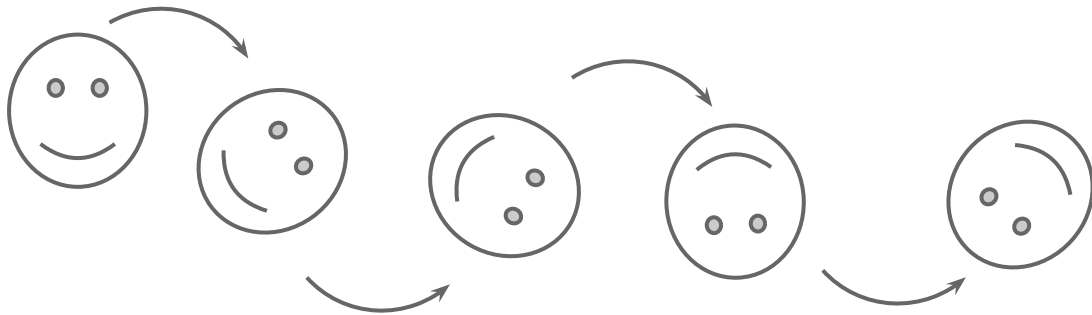
60 days
($2*d$)

30 days
(d)

August 1, 2014
(t)

Did they make **5 edits** during this time?

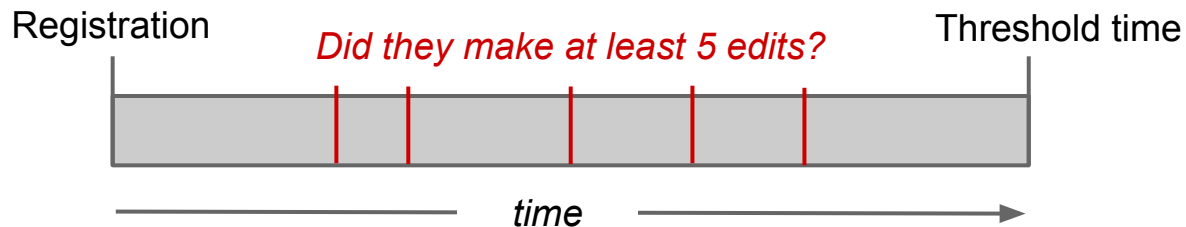
Did they make **5 edits** during this time?





Threshold

Whether editors made a minimum number of edits from the period between registration to the specified threshold hours. Also computes the time it took to reach threshold in hours.

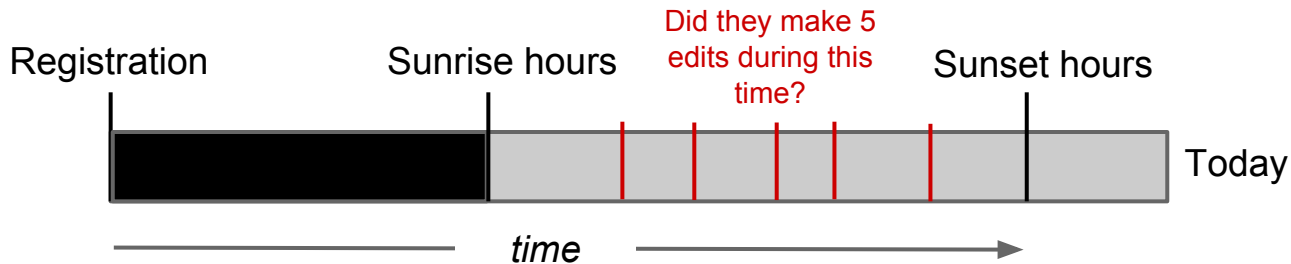


“Edits” = number of times “save page” button is clicked.

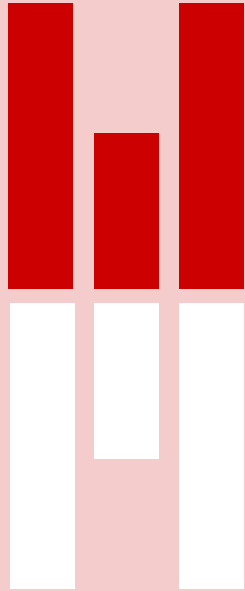


Survival

The number of new users who are making a specified number of edits after a specified number of hours after they register.



"Edits" = number of times "save page" button is clicked.



wikimetrics

demo



Links for the demo

Wikimetrics

metrics.wmflabs.org

Demo cohort

https://en.wikipedia.org/wiki/Wikipedia:Meetup/NYC/NYU_Law

or Use your own cohort



Voluntary Reporting

Open call for program reports!

- Edit-a-thons
- Wikipedia Education Program
- GLAM content donations
- Editing workshops
- Photo contests
- Wiki Loves Monuments
- On-Wiki writing contests
- Conferences
- Hackathons
- Wikimedia in Residence

Wikimetrics IRC chat starting now!

IRC freenode: [#wikimedia-ped](#)

Further support

connect with us:

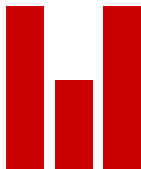
[Evaluation portal News page](#)

email us:

eval@wikimedia.org

talk to us:

[Grants_talk:Evaluation](#)



Resources

Wikimetrics Learning Module - More details!

https://meta.wikimedia.org/wiki/Grants:Evaluation/Learning_modules

Wikimetrics mailing list

wikimetrics@lists.wikimedia.org

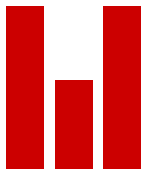
<https://lists.wikimedia.org/mailman/listinfo/wikimetrics>

Phabricator bug report

<https://phabricator.wikimedia.org/maniphest/task/create/?projects=Analytics-Wikimetrics,Analytics-Engineering>

Other questions?

eval@wikimedia.org



Resources

Where to find Wikimetrics and these resources:

meta: [Grants:Evaluation#Measure](https://www.wikimedia.org/wiki/Grants:Evaluation#Measure)

Evaluation portal



Study



Plan



Measure



Share



Connect

Measure your program's success.

Process and Tracking

Tracking and Reporting Toolkit

An easy form to guide you with reporting/tracking information for your program. Translated into 2 languages.

Wikimetrics

A tool for getting user data for your programs

Survey Tools

Survey Question Bank

Qualtrics

Additional resources