目次

[第1章 序論](#_Toc410319897)

[1.1. 本章の構成 1](#_Toc410319898)

[1.2. 研究背景 1](#_Toc410319899)

[1.3. 研究目的 1](#_Toc410319900)

[1.4. 研究方法 2](#_Toc410319901)

[1.5. プロジェクトマネジントとの関連性 2](#_Toc410319902)

[1.6. 成果物のイメージ 2](#_Toc410319903)

[1.7. 本論文の構成 3](#_Toc410319904)

[第2章 Wikiについて](#_Toc410319905)

[2.1. 本章の構成 5](#_Toc410319906)

[2.2. Wikiとは 5](#_Toc410319907)

[2.3. ウィキの文書マークアップ構文について 6](#_Toc410319908)

[2.3.1. Wikiのリンクについて 7](#_Toc410319909)

[2.3.2. Wikiのページ作成について 7](#_Toc410319910)

[2.3.3. Wikiの変更と管理について 7](#_Toc410319911)

[2.3.4. Wikiのユーザ管理について， 8](#_Toc410319912)

[2.4. インターネット百科事典とは 9](#_Toc410319913)

[2.4.1 . インターネット百科事典の特徴 9](#_Toc410319914)

[2.4.2. インターネット百科辞典の種類 9](#_Toc410319915)

[2.4.3. 主なインターネット百科事典 12](#_Toc410319916)

[第3章 Wikipediaについて](#_Toc410319917)

[3.1. 本章の構成 15](#_Toc410319918)

[3.2. Wikipediaとは 15](#_Toc410319919)

[3.3. Wikipediaの新規ページの作成方法 15](#_Toc410319920)

[3.4. ウィキペディア日本語版の記事数の推移 18](#_Toc410319921)

[3.5. 現在ある全言語版のウィキペディアのランキング 19](#_Toc410319922)

[4. 調査・開発](#_Toc410319923)

[4.1. 本章の構成 25](#_Toc410319924)

[4.2. 編集回数の多いウィキペディアンのデータ収集について 25](#_Toc410319925)

[4.2.1. stat.awkで「ファイルの状態」を取得するコマンドを作成する． 25](#_Toc410319926)

[4.2.2. record.awkで「行」のコマンドを作成する． 25](#_Toc410319927)

[4.2.3. pagesize.awkで「ページの大きさ」のコマンドを作成する． 25](#_Toc410319928)

[4.2.4. diff.awkで「テキストファイルの差分を抽出する」コマンドを作成する． 25](#_Toc410319929)

[4.2.5. byte.awkで「バイトの表示をする」コマンドを作成する． 25](#_Toc410319930)

[4.3. Wikipediaの活動調査について 26](#_Toc410319931)

[4.3.1. 利用者ページのサイズ 26](#_Toc410319932)

[4.3.2. 投稿記録 27](#_Toc410319933)

[4.3.3 .編集回数の多いウィキペディアン 27](#_Toc410319934)

[4.3.4. 編集回数の多いウィキペディアンの個人ページサイズ 28](#_Toc410319935)

[4.3.5 . 編集回数の多いウィキペディアンの投稿記録 28](#_Toc410319936)

[4.3.6. 編集回数の多いウィキペディアンのヒストグラム 35](#_Toc410319937)

[5. 調査結果考察](#_Toc410319938)

[5.1. 本章の構成 48](#_Toc410319939)

[5.2. ヒストグラムの結果 48](#_Toc410319940)

[5.3. 考察 62](#_Toc410319941)

[謝辞 63](#_Toc410319942)

表目次

[表 1. Wikiの特徴 5](#_Toc410302345)

[表 2 マークアップ構文とHTMLへの変換例[編集] 6](#_Toc410302346)

[表 3 主なインターネット百科事典 12](#_Toc410302347)

[表 4　ウィキペディア日本語版の記事数の推移 18](#_Toc410302348)

[表 5　100万記事以上ある国 19](#_Toc410302349)

[表 6　10万記事以上100万記事未満の国 19](#_Toc410302350)

[表 7　1万記事以上10万記事未満の国 20](#_Toc410302351)

[表 8　Wikipediaの総合計 23](#_Toc410302352)

[表 9　試験版ウィキペディア 23](#_Toc410302353)

[表 10　Wikipedia：編集回数の多いウィキペディアンランキング 29](#_Toc410302354)

[表 11　最大値と最小値の表 35](#_Toc410302355)

[表 12　編集回数の度数表 42](#_Toc410302356)

[表 13　log10(編集回数の度数)表 42](#_Toc410302357)

[表 14　総編集回数の度数表 43](#_Toc410302358)

[表 15 log10(総編集回数の度数)表 43](#_Toc410302359)

[表 16 個人ページサイズの度数表 44](#_Toc410302360)

[表 17 log10(個人ページサイズの度数)表 44](#_Toc410302361)

[表 18 編集量の度数表 45](#_Toc410302362)

[表 19 log10(編集量の度数)表 45](#_Toc410302363)

[表 20 編集量の標準偏差の度数表 46](#_Toc410302364)

[表 21 log10(編集量の標準偏差の度数)表 46](#_Toc410302365)

[表 22　log10の度数表 53](#_Toc410302366)

図目次

[図 1　ウィキペディア日本語版の記事の推移のグラフ 18](#_Toc410302392)

[図 2　ファイルからオプション 35](#_Toc410302393)

[図 3　オプションからアドイン 36](#_Toc410302394)

[図 4　アドインから設定 36](#_Toc410302395)

[図 5　設定から分析ツール 37](#_Toc410302396)

[図 6　データ分析からヒストグラムを選択 37](#_Toc410302397)

[図 7　入力範囲，データ区間に入力 38](#_Toc410302398)

[図 8　不完全のヒストグラム 38](#_Toc410302399)

[図 9 データ系列の書式設定 39](#_Toc410302400)

[図 10　データ系列の書式設定 39](#_Toc410302401)

[図 11　ヒストグラム 40](#_Toc410302402)

[図 12　COUNTIFSで範囲指定 41](#_Toc410302403)

[図 13 編集回数の度数ヒストグラム 48](#_Toc410302404)

[図 14 総編集回数の度数ヒスとグラム 49](#_Toc410302405)

[図 15 個人ページサイズの度数ヒストグラフ 50](#_Toc410302406)

[図 16　編集量の平均の度数ヒストグラム 51](#_Toc410302407)

[図 17　編集量の標準偏差度数ヒストグラム 52](#_Toc410302408)

[図 18　log10(編集回数)の度数ヒストグラム 54](#_Toc410302409)

[図 19　log10(総編集回数)の度数ヒストグラム 55](#_Toc410302410)

[図 20　log10(個人ページサイズ)の度数ヒストグラム 56](#_Toc410302411)

[図 21 log10(編集量の平均)の度数ヒストグラム 57](#_Toc410302412)

[図 22　log10(編集量の標準偏差)の度数ヒストグラム 58](#_Toc410302413)

[図 23　log10の各度数の折れ線グラフ 59](#_Toc410302414)

[図 24　個人ページサイズと総編集回数の 両対数グラフ 60](#_Toc410302415)

[図 25　個人ページサイズと編集回数の両対数グラフ 61](#_Toc410302416)

第1章

序論

# 序論

* 2. 本章の構成

第１章では，本論文の序論を述べる．研究背景，研究目的，研究方法，プロジェクトマネジメントとの関連性，成果物のイメージ，本論文の構成について述べる．

* 1. 研究背景

Wikipediaは多くのボランティアにより，近年大きな成長をみせたオンライン百科事典プロジェクトである．Wikipedia英語版は2001年1月15日に創設され，日本語版は同年の5月20日に設立した．2003年6月20日に「ウィキメディア財団」がジミー・ウェールズにより創設された．現在2014年9月1日では，287言語のWikipediaが開設されており，この全てのオープンコンテキストの知識資源は無料で一般に提供されている．

多くのボランティアが参加するプロジェクトの代表例であるWikipediaを調査することによって，このような形式のプロジェクトのマネジメントについての有意義な知見が得られることが期待される．

ウィキペディアは，信頼されるフリーなオンライン百科事典，それも質量ともに史上最大の百科事典を，共同作業で創り上げることを目的とするプロジェクト，およびその成果である百科事典本体です[1]

ウィキペディアの成功理由についてはさまざまな考察がされている．例えば，「Google」の検索結果の上位(最初のページ)に「Wikipedia」の項目が表示されていることが多い．これは，Wikipediaにリンクするウェブページがおおくなった結果であるが，このことによって，Wikipediaを利用する人はさらに増えたはずである．

しかし，Wikipediaの衰退が始まっているという報告もある[2]．Wikipediaの英語版では，執筆者の協力を失いつつあり，2007年から2012年までに執筆者の30％を失っている．(現在3万1000人．世界のWikipediaの総計は25万人である[2])．Wikipediaの日本語版の統計[2]を見ても同じような傾向がみられ，2009年11月1日の11,310人に対して，2013年12月12日は，10880人と減少している．

本研究では，参加者の振る舞いのデータをマイニングすることによって，Wikipediaの成功理由を見つけることを目指す．さらに，衰退してきているWikipediaの執筆者の離脱を防ぐ手段について考案する．

* 1. 研究目的

Wikipediaを一つのプロジェクトとみなし，このオンライン百科事典で人的資源がどのように活用されているかを調査する．この調査により，オープンな共同作業プロジェクトにおける，人的資源マネジメントのあり方についての知見を得たい．

* 1. 研究方法

Wikipediaの作業履歴データを取得し，どのように人的資源が活用されているかを明らかにする．編集者の作業履歴データを取得後，グラフを作成し，編集者にどのような傾向がみられるか調査する．また，オープンな共同作業プロジェクトにおける人的資源マネジメントのあり方を提案する．

* 1. プロジェクトマネジントとの関連性

　Wikipediaプロジェクトでは，PMBOKに記されている，人的資源マネジメントについて，関連性が最もあるものである．記事の作成や，変更修正には知識を持った人々の協力が必要不可欠であり，オープンな共同作業でWikipediaを創り上げ，その成果であるものが，この百科事典本体である．

プロジェクト人的資源マネジメントは，プロジェクト・チームを組織し，マネジメントし，リードするためのプロセスからなる．プロジェクト・チームは．プロジェクトを完了するための役割と責任を割り当てられた人々で構成される．プロジェクト・チーム・メンバーの種類と数は，プロジェクトの進行と共にしばしば変更する．プロジェクト・チーム・メンバーは．プロジェクトのスタッフと呼ばれることもある．プロジェクト・チーム・メンバーには特定の役割と責任を割り当てられるが．プロジェクトの計画と意思決定にはチーム・メンバー全員が関与することが望ましい．初期の段階からチーム・メンバーが関与し，参加することにより，計画策定のプロセスにおいて，専門知識が取り込まれ，かつプロジェクトへの参加意欲が高まる[3]

プロジェクトマネジメント・チームは，プロジェクト・チームの一部であり，さまざまなプロジェクト・フェーズの立ち上げ，計画，実行，監視，コントロール，終結等のプロジェクトマネジメント活動やリーダーシップ活動を担当する，このグループは，コア・チーム，幹部チーム，または統率チームと呼ばれることもある．小規模なプロジェクトでは，プロジェクトマネジメントの責任をチーム全体で共有することも，プロジェクト・マネジャー1人で担うこともある．プロジェクト・スポンサーはプロジェクトマネジメント・チームに協力し，一般には，プロジェクトの資金調達，スコープの明確化，進捗の監視などを支援し，プロジェクトの利益になるように他の人々に働きかける[3]

* 1. 成果物のイメージ

Wikipediaで行われている人的資源がどのように活用されていているかを調査し傾向をグラフにまとめる．さらに，オンラインのオープンな共同作業プロジェクトでの人的資源マネジメントの知見を得る．

* 1. 本論文の構成

第1章では，序論を記述する．第2章では，本論文の研究目的である，Wikipediaについて，Wikipediaがどのようなものなのか，どのように記事が作成されるのかを調査する．さらに，記事の修正変更はどのように行われるのかなどを記述する．第3章では，Wikipediaの作業履歴データの取得を解説し，人的資源マネジメントがどのように活用されているのか調査する．第4章では，プロジェクトマネジメントとの関係を記述する．第5章では，具体的な調査の方法や過程を記述する．第6章では，本研究についての調査結果・考察，まとめを記述する．

参考文献

[1]Cave cattum (2008) 「Wikipedia:ウィキペディアについて」<http://ja.wikipedia.org/wiki/Wikipedia:%E3%82%A6%E3%82%A3%E3%82%AD%E3%83%9A%E3%83%87%E3%82%A3%E3%82%A2%E3%81%AB%E3%81%A4%E3%81%84%E3%81%A6> (参照2013-7-3)

[2]Alessandro Longo (2013) 「Wikipediaが岐路に立っている？」 <http://wired.jp/2013/11/27/the-decline-of-wikipedia/> (参照2013-11-27)

[3] Project Management Institute, Inc. プロジェクトマネジメント知識体系ガイド(第4版). PMI, 2009-12.

# 第2章 Wikiについて

* 2. 本章の構成

本章では，本論文の研究対象であるWikipediaについて，Wikiがどのようなものなのか，基本知識や使われ方を述べる．また，インターネット百科事典とはどのようなものか，インターネットでの共同作業について調査し，Wikipediaではどのように記事が作成されるのかを調査し述べる．

* 1. Wikiとは

Wiki(ウィキ)とは，ウェブブラウザを利用し，Webサーバ上でのハイパーテキスト文書を書き換えるためのシステムの1種であるWikiの使用用途では普通，ネットワーク上で誰でも，どこからでも，文書の書き換えができるようになっており，オンラインなら共同作業で文書の作成などができる．Wikiは共同作業で文書を作成できる特徴から，コラボレーションツールやグループウェアであるとも評されている．ソフトウェアとしては，始めに作られたプログラムに改良や改変を行え，またそのプログラムを参考にし，今日ではWikiが数多く活用されている．

以下の表1.に，多くのWikiに共通している特徴を記す．

表 1. Wikiの特徴

|  |
| --- |
| ・ネットワーク上であれば，誰がいつ，どこからでも文書の書き換えを行うことができる． |
| ・文書を書き換えるために最低限必要なツールは，ウェブブラウザである． |
| ・Wikiの特有な文書マークアップは，HTMLなどと比べると，簡潔にできているので覚えやすい． |
| ・同じWiki内での文書間にリンクが貼りやすくなっており，個々の文書を高度に連携した文書郡が作成しやすい． |
| ・大抵は，記事の変更の事前許可を執拗としておらず，Wikiのあるサーバを接続できる人に開かれている． |
| ・ユーザアカウントの登録を必要としていないところも多くある． |

* 1. ウィキの文書マークアップ構文について

大抵のウィキはそれぞれ独自のマークアップ構文を策定していて，ウィキが扱う文書はその構文に従って記述され，そのまま文書ファイルとして保存される．そして文書が閲覧されるときには，文書内に記述されたマークアップはウィキプログラムによって適切なHTML形式に変換されて，ウェブブラウザはその変換されたHTML文書を表示することになる．

表 2 マークアップ構文とHTMLへの変換例[編集] 引用[4]

|  |  |  |
| --- | --- | --- |
| **ウィキ文** | **変換後のHTML** | **表示例** |
| [[文部科学省|文科省]] | **<A** href="/wiki/文部科学省" title="文科省"**>**文科省**</A>** | 文科省 |
| \*国語辞典 \*漢和辞典 \*類義語辞典 | **<UL><LI>**国語辞典**</LI><LI>**漢和辞典**</LI><LI>**類義語辞典**</LI></UL>** | * 国語辞典 * 漢和辞典 * 類義語辞典 |
| 鳴かぬなら，''鳴くまで待とう''，ホトトギス | 鳴かぬなら，**<I>**鳴くまで待とう**</I>**，ホトトギス | 鳴かぬなら，*鳴くまで待とう*，ホトトギス |
| 鳴かぬなら，'''鳴くまで待とう'''，ホトトギス | 鳴かぬなら，**<STRONG>**鳴くまで待とう**</STRONG>**，ホトトギス | 鳴かぬなら，**鳴くまで待とう**，ホトトギス |

上の例では，HTMLでは <strong>〜</strong> タグを使って強調するところを，ウィキ構文では '''〜''' を使っていたり，ウィキ構文の二重スクエア・ブラケットで単に ウェブブラウザと書いたところが，同じウィキ内の別の文書へのリンクに変換されているのが見て取れる．

HTML自体は高機能で，豊富な種類の要素を複雑に入れ子にしたり，見栄えを調整するスタイルシートなどを埋め込んだりすることもできる．一般的なウィキにおいては，むしろこれらの機能を制限することによって，文書の作成・編集を容易にするとともに，個々のユーザーがスタイルを埋め込んでしまい全体としての一貫性が崩れるという危険性を回避している．また同時に，文書の見栄えではなく内容の方へ利用者の注意を集中させるという効果も狙っている．

それでも，現在のウィキエンジンでは記事の編集にActiveXコントロールや Ajax などで作られた WYSIWYG エディタを内蔵しているものもある．これにより，クライアントサイドでの編集を容易にすることができる．ブラウザ外部のフロントエンドツールを用いて編集を簡素化するツールも存在する．

ウィキのマークアップ構文は使用するウィキエンジンごとに様々である．簡単なウィキでは基本的なテキストのフォーマットのみが用意され，もっと複雑なウィキでは，表，画像，テンプレートによる定型文など，さらには投票やゲームまで実現するものもある．ただ，あまりに多様なので，標準化しようという動きがある引用[4]．

### Wikiのリンクについて

Wikiはハイパーテキストメディアで，読者の検索ワードに関係するページを容易にナビゲーションする構造を持っている．各ページに他のページへの多数のリンクを持っている．大きなWikiには，ハイパーテキストメディアで階層的なナビゲーションが存在しているが，必ずしも使われる必要はない．リンクは特定の構文(リンクパターン)を用いて作成される．

* + 1. Wikiのページ作成について

Wikiで新しくページを作成するには，厳密には他のページとリンクされて作成されなければならない．リンクは，トピックとの関係のあるページに作成される．リンクが存在しなければ，そのリンクには「壊れた」リンクとして表示される．そのリンクを開くとページを編集するページが開き，ユーザはそのページに新しくテキストを入力することができる．このような仕組みによって，リンクの全くされていない「孤立したページ」が作成される割合が格段に下げられ，高い関連性のあるページ群が保持されることになる．

### Wikiの変更と管理について

Wikiでは一般に「間違いをおかしにくくするのではなく間違いを直しやすくする」という哲学に従って作られている．Wikiでは，そのために非常にオープンである一方で，ページ内容の変更に妥当性を検証するための手法を持っている．

Wikiの殆どすべてにある最もすぐれている機能は，「最近更新したページ」である．これは，最近編集されたものに番号を付けた詳細なリストや，決まった期間に編集が行われたすべてのリストである．いくつかのWikiでは，少しの編集や自動インポートスクリプトで行われる編集をフィルタして表示させないようにすることも可能である．

　Wikiの大半は，更新ログから２つの機能を利用することができる．１つは「改訂履歴」であり，ページの以前の版をみられることである．もう１つは，「差別機能」であり，２つの版の差異を強調して表示できるものである．改訂履歴の機能を使うと，以前の版のページを開くことや，保存することもでき，このことにより，変更される以前の内容へと復元することも可能である．差異機能は，Wikiの利用者が最近のページ更新にリストされた差異をみて，許容できない編集がされていた場合，それを以前のものに戻す必要があるか判断するために使うことができる．この手順は，利用しているWikiにもよるが，多かれ少なかれ自動化がされている．改訂履歴を保存し，過去の任意のページの版に戻す，revert機能を提供しているWikiエンジンも多くある．

* + 1. Wikiのユーザ管理について，

多くのWikiでは，ユーザ登録を義務づけすることは避けているが，事実上ではすべての大きなWikiエンジンは，コミュニティのルールを常習的に違反するユーザを制限するためのいくつかの方法を備えている．もっとも一般的な方法では，ある特定のユーザの編集行為を禁止することである．これは，特定のIPアドレスからのアクセスを禁止することである．しかし，多くのインターネットサービスプロバイダでは，接続のたびに新しいIPアドレスを割り振ってしまうので，IPアドレスを用いたアクセス制限には，比較的簡単にすり抜けられてしまっている．また，IPアドレスを用いたアクセス制限では，無関係なユーザのアクセスを請願してしまう結果になってしまうこともある．

小さなWikiでは，常習的にルールを破るユーザへの共通の防御手段は，ルールを破るユーザたちに，ページを好きなだけ荒らさせて，荒らしが去った後にそのページをすぐに復旧させることである．この防御手段は，大きなコミュニティの状況では，しばしば受け入れられにくいと考えられる．もっと素早いアクションが好まれるからである．変動IPアドレスの問題を処理するには，時限式の編集禁止措置を行い，特定の範囲すべてのIPアドレスの禁止へと拡げられる場合もある．これが抑止力として十分である場合が多いということを背景としていて，これにより，常習的にルールを破るユーザがある期間内だと編集ができなくすることが可能になる．

緊急措置として，いくつかのWikiではデータベースを読み出ししかできないように切り替えることができる．または，期日までに登録されたユーザにだけ，編集を続けさせることができる．しかし，一般的にいうと，常習的にルールを破るユーザによるどんな損傷でも素早く復旧させることができる．問題なのは，微妙な間違いがページの中に紛れ込み，他の人が気が付かなくなっていくことである．

多くのWikiでは，ある特定のページへの編集を凍結させることができる機能を備えている．ほとんどのWikiにおいて，この機能を使うのは極端な場合のみに限られ，滅多なことでは使われることはない．Wikipediaの場合，ページの「保護」と呼ばれており，保護されたページは保護を行使したり，解除したりできる管理者権限を持つ，一部の人しか編集できない．これを使うことは，一般にWikiの基本哲学に反していると考えられるので，可能ならばできる限り避けるべきであるとされている．

* + 1. Wikiの検索について

Wikiでは大抵，全文検索ではなくとも，少なくとも記事の名前検索を提供している．検索の拡張性は，Wikiエンジンがデータベースを使っているか，使っていないかに強く依存している．データベースの検索呼び出し機能は大きなWikiでの高速検索には必須である．Wikipediaでは，表示ボタンで読者が検索条件のキーワードを入力して，そのキーワードにできるだけ合致するページを，直接見つけられるようになっている．いくつかのWikiを同時進行で検索するために，メタWiki検索エンジンが作られた．

* 1. インターネット百科事典とは

　インターネットでアクセスができるデジタルの百科事典である．インターネット百科事典として今日，最も規模の大きいものがWikipediaである．また，専門家が主導しているオープンアクセスのプロジェクトで歴史が古くからあるものとして，スタンフォード哲学百科事典などがある．

### 2.4.1 . インターネット百科事典の特徴

インターネット百科事典には，旧来の書籍型の百科事典と比べていくつかの違いがある．例えば挙げられる利点として次のようものがある[29][30][31][15]．

まずデジタルであることの利点として，テキスト検索が行える事，保存に必要な物理的スペースが書籍に比べて小さいこと，技術の進歩に合わせて動画やプログラムなど新しい情報の提示方法を取り込んでいける事，などがある．またインターネットを使用していることから来る利点として，改訂スピードの早さ，製作コストの低さ，ネット環境さえあればどこからでもアクセスできる利便性，ページ数という制約からの解放，ウェブ上の他の情報リソースに直接リンクできること，などがある．

問題点としては，上記の利点の裏返しであるが，次のような点がある．完成した項目から公開されていくため，事典全体としては歯抜け，つまり未完成の状態となりやすいこと．ネットにアクセスできない環境で利用しにくいこと，パソコンの状態の不調などの技術的な理由で使えなくなる場合があること，購入した書籍の百科事典と違い運営元が破綻すると閲覧できなくなってしまうこと，過去版を保存していないサイトでは引用や参照がおかしくなる場合があること，などがある．またウィキペディアのような誰もが執筆できるプロジェクトではその情報の質・信頼性がしばしば問題となる[32][33][34]．引用[5]

### 2.4.2. インターネット百科辞典の種類

　インターネット百科事典には，使用されている言語の違いや，取り扱っている主題の違い，運営方法まで，様々な違いがある．

オープンアクセス/購読

　ウィキペディアやコロンビア百科事典，スカラーペディアなどは，インターネット経由で誰もが無料で閲覧できるスタイルを採用している．これはオープンアクセスモデルなどと言われる[35]．

一方で，ブリタニカ・オンラインやラウトレッジ哲学百科事典などは，一定期間ごとに決められた金額を支払った者のみが全文を閲覧できる，購読型のモデルを採用している．

この場合，個人で料金を支払う方法を個人購読（Personal Subscription），大学図書館や研究所などが組織単位で料金を支払う契約方法を機関購読（Institutional Subscription）と言う．機関購読で契約した場合，特定のIPアドレスからのアクセスに対して閲覧を許可する，といった方法が取られる．引用[5]

ライセンス

百科事典の内容を再利用できるかどうかで，サイトごとに様々な違いがある．各国の著作権法や条約で定められているすべての権利を主張するAll rights reservedの状態で公開されているものと，クリエイティブ・コモンズ，GFDLといった再利用可能なライセンスの元で公開されているものがある．

例えばウィキペディアではCC-BY-SAというライセンスが採用されており，著作者表示など一定の用件を満たしさえすれば，サイトの内容を書籍の形で有償で販売することも自由に可能である．一定の条件の元で自由に再利用が可能であることはオープンコンテントとも言われる[36]．引用[5]

執筆者

専門家が執筆する百科事典では，編集委員が各分野の専門家に依頼して各項目を執筆してもらうのが一般的である．珍しい例として，スカラーペディアには利用者の投票で執筆者を選定する選挙システムがある．

原稿を執筆した専門家に対して謝礼を支払うかどうかはサイトによって異なる．ブリタニカは執筆者に金銭を支払っているが，スカラーペディアやスタンフォード哲学百科事典は支払っていない[40]．

ウィキペディアなどの百科事典サイトでは執筆者は自由参加となっており，無休でかつ参加が痔発性に基づく事からボランティアと表現されることもある[41][42]．こうしたサイトではコンテンツの利用者と作成者は近い位置にあるか，または重なっている．こうした作り手と受けてが重なった状態にあることはWeb 2.0，またそうして生み出される媒体やコンテンツはConsumer Generated Media（CGM，消費者生成メディア），User Generated Content（UGC，ユーザー生成コンテンツ）などと言われる[43][44][45]．

匿名で参加できる自由参加型プロジェクトにおける投稿者像はそれほどはっきりしない．2009年にウィキペディアで13万人を対象に行われたアンケート調査では，投稿者の75%が30歳以下，そして投稿者の87%が男性であった[37][38][39]．引用[5]

資金源

運営資金の源はサイトにより様々である．購読型のモデルで運営されているサイトは購読料収入がある．広告の掲載されているサイトには広告料収入がある．また内容と関連する企業・団体がスポンサーとして付く場合もある．一定の公益性のあるサイトであれば，寄付や，各種財団・大学・政府機関などからのグラントが資金源となり得る．

資金内容の公開状況はサイトによってマチマチである．以下，期間やデータは統一されたものでないが，各プロジェクトで公表されている予算関連のデータをいくつかあげる．引用[5]

・ウィキペディア - 2010年度予算（1年間） 収入および支出 2040万ドル（主に寄付）[46]

・スタンフォード哲学百科事典 - サイト設立の1995年から10年間で計約200万ドルのグラントを取得．2007年から先3年間で，様々な図書館からの資金提供で300万ドル，個々人の寄付金で112万5000ドル，合計412万5000ドルを集める予定（2009年12月時点で75%達成済み）[40][47]

・Citizendium - 2011年2月度予算（1ヶ月間） 収入48ドル（寄付），支出 320ドル（サーバーホスティング）[48][49]

・Encyclopedia of Life - 2008年の設立から10年以内に地球上の動植物180万種 全てについての記事を完成させることを目指すプロジェクト[50]．全体で必要な予算は計1億1050万ドル程と見積もられている[51]．この内，2007年5月時点でマッカーサー基金が1000万ドル，スローン財団が250万ドル，計1250万ドルの補助金を提供する事をすでに決定している[52]

・ブリタニカ - 18世紀半ばに創刊した同百科事典は，1989年には書籍版の売り上げが年間6億5000万ドルあった．その後CD-ROM化とオンライン化の波の中でマイクロソフト社のエンカルタに押され経営危機に陥り，1996年にスイスの投資会社ジェイコブ・サフラ財団に売却された．オンライン版は1994年に有料サービスとしてスタート．1999年に全面無料化するも，2001年には再び有料に戻す．2009年には読者による編集を可能にするなど，経営方法の模索が続いている[15][53]引用[5]

### 2.4.3. 主なインターネット百科事典

表 3 主なインターネット百科事典 引用[5]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **名称** | **開始年** | **運営者** | **言語** | **ソース** |
| ATOMICA | 1996年10月2日 （インターネット版） | 高度情報科学技術研究機構（RIST） | 日本語 | 専門家に執筆を依頼 |
| ウィキペディア | 2001年1月15日 | ウィキメディア財団 | 多語 | 参加型（資格不要） |
| EnCyber | 2003年 | NHN株式会社 | 韓国語 | 斗山世界大百科事典 |
| コロンビア電子百科事典 | 2004年 | コロンビア大学 | 英語 | コロンビア百科事典 |
| アンサイクロペディア | 2005年1月 | アンサイクロペディア財団 | 多言語 | 参加型（資格不要） |
| Answers.com | Answers Corporation | 英語 | 他のウェブサイト |
| Weblio | 2005年12月12日 | ウェブリオ株式会社 | 日本語 |
| 百度百科 | 2006年4月20日 | 百度公司 | 中国語 | 参加型 |
| ユアペディア | 2007年4月4日 | Beyond（吉本敏洋） | 日本語 | 参加型（資格不要） |
| ニコニコ大百科 | 2008年5月12日 | 有限会社未来検索ブラジル | 参加型（ニコニコ動画プレミアム会員） |
| はてなキーワード | 2008年9月1日 | 株式会社はてな | 参加型（はてなダイアリー市民） |
| Yahoo!百科事典（ベータ版） | 2008年11月26日 | ヤフー株式会社 | 日本大百科全書（ニッポニカ） |
| ウィークペディア | 2009年3月20日 | Hanakayu | 参加型（資格不要） |
| エンペディア | 2009年4月21日 （2013年1月1日に改めて開始） | エンペディアコミュニティとスタッフ |
| コトバンク | 2009年4月23日 | 朝日新聞，講談社，小学館，朝日新聞出版，ECナビ | マイペディア，知恵蔵，デジタル大辞泉など |

参考文献

[4] Yoshitaka Mieda (2003) 「ウィキ」<http://ja.wikipedia.org/wiki/%E3%82%A6%E3%82%A3%E3%82%AD> (参照2014-10-14)

[5]Kappas (2009) 「インターネット百科事典」<http://ja.wikipedia.org/wiki/%E3%82%A4%E3%83%B3%E3%82%BF%E3%83%BC%E3%83%8D%E3%83%83%E3%83%88%E7%99%BE%E7%A7%91%E4%BA%8B%E5%85%B8> (参照2014-10-15)

# 第3章 Wikipediaについて

* 2. 本章の構成

　本章では，Wikipediaがどのようにできているか調査し，Wikipediaの作業履歴データの取得を解説する．また，Wikipediaにおいて人的資源マネジメントがどのように活用されているのかを述べる．

* 1. Wikipediaとは

　Wikipediaはインターネット上で見ることのできるフリーな百科事典である．Wikipediaは非営利団体であるが，今日では最も大きい百科事典となっている．この百科事典は不特定多数の人々によって記事が作成され，共同で百科事典を作成することをも汽笛としているプロジェクトである．ウィキペディアは2001年1月15日にジミー・ウェールズの個人的なプロジェクトとして立ち上げられた．2003年6月以降からはウィキメディア財団という非営利団体により運営されている．非営利団体なので資金面では多くの方々からの寄付によって運営されている．

今日の2015年1月16日では，288言語のウィキペディアが開設されている．しかし，そのうちの11言語は執筆者が集まらないなどの為に凍結されている言語も含まれている．全言語版の総計では2,500万以上の記事が作成されており，2015年1月1日では，日本語版の記事数は940,107件が作成されている．

* 1. Wikipediaの新規ページの作成方法

新規ページを作成する方法は，大きく5つに分類できます．通常の新規記事を作成したい場合は1〜5のいずれかで，リダイレクトページを作成したい場合は2〜4のいずれかで行ってください．

1.既存のページから作成する．

2.プレビューを利用して作成する．

3.検索結果画面を利用して作成する．

4.ブラウザにURLを直接入力する．

5.既に作成されている下書きページを移動することで作成する．

引用[7]

新規ページを作成する方法

この方法は，執筆したい記事名がすでに記事内に含まれている場合に用いられるものです．たとえば，Aという名の新しい記事を執筆したい場合に，Bという既存の記事にAという言葉が登場しているときに利用します．引用[7]

記事名にしたい言葉にリンクが張ってある場合

1. 執筆したい記事名が含まれている記事を表示する（B記事を表示）．
2. 赤リンク（未執筆[リンク](http://ja.wikipedia.org/wiki/Help:%E3%83%AA%E3%83%B3%E3%82%AF)）をクリックして編集画面にする（Aをクリック）．

引用[7]

記事名にしたい言葉にリンクが張っていない場合

1. 執筆したい記事名が含まれている記事を表示する（B記事を表示）．
2. 「編集」をクリックして編集画面にし，記事名にしたい言葉を[[ ]]で囲む（例:[[A]]）．
3. 「編集内容の要約」に「lk」と入力する（"link"の略）．
4. 「以上の記述を完全に理解し同意した上で投稿する」ボタンをクリックする（これで赤リンクができる）．
5. 赤リンクをクリックして編集画面にする（Aをクリック）

引用[7]

プレビューを利用して作成する方法

既存のページに手を加えずに新しいページを作る方法のひとつに，プレビューを利用する方法があります．手順は次のとおりです．

1. 適当にページを表示する（たとえば，サイドバーにある「練習用ページ」のクリックで開かれる[Wikipedia:サンドボックス](http://ja.wikipedia.org/wiki/Wikipedia:%E3%82%B5%E3%83%B3%E3%83%89%E3%83%9C%E3%83%83%E3%82%AF%E3%82%B9)のページなど）．
2. 「編集」をクリックする．
3. テキストボックスに自分の作りたい記事名のリンクを書く（例:[[新規項目]]）．
4. 「プレビュー」を押す．
5. すると，記事名が赤リンクになるので，クリックする．
6. 新規ページ編集画面が現れる．

引用[7]

検索結果画面を利用して作成する方法

[検索](http://ja.wikipedia.org/wiki/Help:%E6%A4%9C%E7%B4%A2)結果画面を利用する方法です。まず、作成したい記事名をサイドバーの検索ボックスに入力して「表示」または「検索」をクリックします。つぎに、「検索結果」ページ上部の「**このウィキでページ"<記事名>"を新規作成する**」の**<記事名>**の部分（赤色のリンク）をクリックします。すると検索語を記事名とした記事の編集画面になります。引用[7]

ブラウザにURLを直接入力する方法

既存のページに手を加えずに新しいページを作る場合には、ブラウザの[URL](http://ja.wikipedia.org/wiki/Uniform_Resource_Identifier)入力欄に直接入力するほうが簡単です。次のように入力してください。

http://ja.wikipedia.org/wiki/記事名

すると、[このように](http://ja.wikipedia.org/wiki/MediaWiki:Noarticletext)表示されますので、そこから「編集」をクリックします。

**注意**：編集を開始する前に項目名の部分が[文字化け](http://ja.wikipedia.org/wiki/%E6%96%87%E5%AD%97%E5%8C%96%E3%81%91)せずに表示されているか確認してください。ブラウザの種類によっては日本語が文字化けします。そのようなブラウザをお使いの方は、URLを直接入力する方法は使用できませんので、他の方法で作成してください。

下書きページの移動

「新規ページを作成する方法」節のいずれかの方法で下書きを作成してください。下書きを作成する目的で、[利用者ページ](http://ja.wikipedia.org/wiki/Wikipedia:%E5%88%A9%E7%94%A8%E8%80%85%E3%83%9A%E3%83%BC%E3%82%B8#.E5.88.A9.E7.94.A8.E8.80.85.E3.83.9A.E3.83.BC.E3.82.B8.E3.81.A7.E3.81.A7.E3.81.8D.E3.82.8B.E3.81.93.E3.81.A8)を用いることができます。その後、下書きを目的の記事名に[移動](http://ja.wikipedia.org/wiki/Help:%E3%83%9A%E3%83%BC%E3%82%B8%E3%81%AE%E7%A7%BB%E5%8B%95)することによって、記事を作成します。引用[7]

新規ページの執筆と保存の方法は以下の1～4の手順で行える．

1. 説明文の下にあるテキストボックスに内容を記述し，そして編集内容の要約欄にもようやくを記述する．
2. 「プレビュー」を押してどのように表示されるか確認する．
3. 修正が必要であれば，テキストボックスの中身を手直しして，さらに「プレビュー」で確認する．
4. 問題がなければ「以上の記述を完全に理解し同意した上で投稿する」を押す．

引用[7]

## ウィキペディア日本語版の記事数の推移

表 4　ウィキペディア日本語版の記事数の推移 引用[8]



ウィキペディア日本語版の記事の推移のグラフを作成した．

図 1　ウィキペディア日本語版の記事の推移のグラフ

2002年9月～2015年1月の期間でグラフを作成し，縦軸を記事数，横軸を月で記したものである．

## 現在ある全言語版のウィキペディアのランキング

表 5　100万記事以上ある国

|  |  |  |  |
| --- | --- | --- | --- |
| 順位 | 言語 | 純記事数 | 総項目数 |
| 1 | 英語 | 4,695,840 | 34,813,920 |
| 2 | スウェーデン語 | 1,953,022 | 4,524,559 |
| 3 | オランダ語 | 1,807,045 | 3,302,105 |
| 4 | ドイツ語 | 1,800,109 | 4,953,243 |
| 5 | フランス語 | 1,580,812 | 7,093,408 |
| 6 | ワライ語 | 1,259,031 | 2,864,394 |
| 7 | セブアノ語 | 1,205,518 | 2,783,593 |
| 8 | ロシア語 | 1,179,545 | 4,008,319 |
| 9 | イタリア語 | 1,169,434 | 3,902,475 |
| 10 | スペイン語 | 1,150,306 | 4,771,758 |
| 11 | ベトナム語 | 1,111,945 | 3,016,451 |
| 12 | ポーランド語 | 1,086,827 | 2,199,271 |

表 6　10万記事以上100万記事未満の国

|  |  |  |  |
| --- | --- | --- | --- |
| 順位 | 言語 | 純記事数 | 総項目数 |
| 13 | 日本語 | 942,481 | 2,665,095 |
| 14 | ポルトガル語 | 862,116 | 3,700,563 |
| 15 | 中国語 | 806,223 | 3,859,033 |
| 16 | ウクライナ語 | 547,873 | 1,613,145 |
| 17 | カタルーニャ語 | 447,431 | 1,125,371 |
| 18 | ペルシア語 | 440,941 | 2,899,366 |
| 19 | ノルウェー語（ブークモール） | 404,538 | 1,032,331 |
| 20 | フィンランド語 | 364,214 | 970,061 |
| 21 | インドネシア語 | 353,654 | 1,496,724 |
| 22 | アラビア語 | 345,361 | 1,993,584 |
| 23 | セルビア語 | 313,559 | 1,122,630 |
| 24 | チェコ語 | 312,511 | 806,145 |
| 25 | 朝鮮語 | 300,923 | 1,014,453 |
| 26 | ハンガリー語 | 276,369 | 927,044 |
| 27 | セルビア・クロアチア語 | 276,253 | 3,517,229 |
| 28 | マレー語 | 271,861 | 707,070 |
| 29 | ルーマニア語 | 257,578 | 1,146,609 |
| 30 | トルコ語 | 238,675 | 1,176,253 |
| 31 | ミナンカバウ語 | 221,165 | 227,874 |
| 32 | カザフ語 | 210,046 | 492,771 |
| 33 | エスペラント | 208,496 | 455,358 |
| 34 | バスク語 | 205,418 | 553,633 |
| 35 | スロバキア語 | 198,093 | 425,230 |
| 36 | デンマーク語 | 194,887 | 651,154 |
| 37 | ブルガリア語 | 187,558 | 399,835 |
| 38 | リトアニア語 | 169,582 | 376,403 |
| 39 | ヘブライ語 | 165,205 | 696,958 |
| 40 | クロアチア語 | 151,005 | 414,812 |
| 41 | アルメニア語 | 144,096 | 494,491 |
| 42 | スロベニア語 | 143,903 | 322,762 |
| 43 | エストニア語 | 129,950 | 331,995 |
| 44 | ウズベク語 | 127,831 | 644,596 |
| 45 | ヴォラピュク | 119,540 | 248,255 |
| 46 | シンプル英語 | 119,442 | 344,050 |
| 47 | ガリシア語 | 119,390 | 265,069 |
| 48 | ノルウェー語（ニーノシュク） | 118,578 | 272,290 |
| 49 | ヒンディー語 | 117,045 | 497,538 |
| 50 | ギリシア語 | 114,159 | 296,679 |
| 51 | ラテン語 | 112,457 | 210,752 |
| 52 | アゼルバイジャン語 | 100,672 | 253,784 |

表 7　1万記事以上10万記事未満の国

|  |  |  |  |
| --- | --- | --- | --- |
| 順位 | 言語 | 純記事数 | 総項目数 |
| 53 | タイ語 | 91,977 | 517,916 |
| 54 | オック語 | 88,222 | 141,033 |
| 55 | グルジア語 | 87,822 | 249,654 |
| 56 | マケドニア語 | 80,334 | 1,075,875 |
| 57 | ベラルーシ語 | 79,464 | 193,871 |
| 58 | チェチェン語 | 78,559 | 92,446 |
| 59 | ネパール・バサ語 | 71,164 | 193,705 |
| 60 | タミル語 | 66,215 | 197,849 |
| 61 | タタール語 | 65,974 | 148,159 |
| 62 | ウルドゥー語 | 64,391 | 259,738 |
| 63 | ピエモンテ語 | 63,696 | 93,552 |
| 64 | タガログ語 | 63,197 | 202,822 |
| 65 | ウェールズ語 | 61,228 | 143,215 |
| 66 | テルグ語 | 60,258 | 165,430 |
| 67 | ラトビア語 | 58,237 | 237,822 |
| 68 | ボスニア語 | 56,736 | 314,681 |
| 69 | ベラルーシ語（タラシケヴィツァ） | 55,260 | 143,095 |
| 70 | マダガスカル語 | 53,382 | 185,614 |
| 71 | ハイチ語 | 52,967 | 61,449 |
| 72 | ブルトン語 | 52,192 | 106,639 |
| 73 | アルバニア語 | 52,065 | 127,162 |
| 74 | ジャワ語 | 48,370 | 103,573 |
| 75 | ルクセンブルク語 | 42,687 | 95,100 |
| 76 | マラーティー語 | 41,103 | 153,468 |
| 77 | アイスランド語 | 38,543 | 104,907 |
| 78 | マラヤーラム語 | 37,766 | 258,821 |
| 79 | 広東語 | 34,332 | 97,294 |
| 80 | アフリカーンス語 | 34,120 | 83,539 |
| 81 | ベンガル語 | 33,762 | 368,338 |
| 82 | バシキール語 | 33,445 | 73,525 |
| 83 | 西パンジャーブ語 | 33,243 | 40,244 |
| 84 | アイルランド語 | 32,662 | 54,714 |
| 85 | ビルマ語 | 32,589 | 49,077 |
| 86 | ロンバルド語 | 31,509 | 81,294 |
| 87 | 西フリジア語 | 31,347 | 71,663 |
| 88 | ヨルバ語 | 31,084 | 53,375 |
| 89 | タジク語 | 30,570 | 57,517 |
| 90 | アラゴン語 | 30,415 | 98,800 |
| 91 | チュヴァシ語 | 30,342 | 49,436 |
| 92 | スコットランド語 | 29,208 | 82,126 |
| 93 | スワヒリ語 | 27,768 | 70,893 |
| 94 | キルギス語 | 27,472 | 42,826 |
| 95 | ネパール語 | 26,969 | 67,266 |
| 96 | イド語 | 26,230 | 38,336 |
| 97 | グジャラート語 | 25,645 | 53,816 |
| 98 | ビシュヌプリヤ・マニプリ語 | 25,125 | 57,600 |
| 99 | シチリア語 | 24,741 | 54,993 |
| 100 | 低ザクセン語 | 22,437 | 55,998 |
| 101 | クルド語 | 20,893 | 50,121 |
| 102 | アストゥリアス語 | 20,067 | 39,032 |
| 103 | ケチュア語 | 19,237 | 50,893 |
| 104 | スコットランド・ゲール語 | 19,205 | 30,286 |
| 105 | アレマン語 | 17,988 | 42,002 |
| 106 | スンダ語 | 17,807 | 38,093 |
| 107 | カンナダ語 | 17,475 | 58,443 |
| 108 | パンジャーブ語 | 16,404 | 50,088 |
| 109 | アムハラ語 | 16,252 | 42,005 |
| 110 | ソラニ語 | 15,147 | 94,557 |
| 111 | インターリングア | 14,352 | 24,607 |
| 112 | ナポリ語 | 14,330 | 28,326 |
| 113 | ブギス語 | 14,090 | 18,239 |
| 114 | モンゴル語 | 14,012 | 38,681 |
| 115 | サモギティア語 | 13,765 | 23,428 |
| 116 | ワロン語 | 13,680 | 35,107 |
| 117 | アラビア語エジプト方言 | 13,305 | 93,283 |
| 118 | バニュマス語 | 13,048 | 26,072 |
| 119 | シンハラ語 | 12,214 | 49,380 |
| 120 | マーザンダラーン語 | 12,038 | 24,070 |
| 121 | 閩南語 | 12,011 | 29,927 |
| 122 | イディッシュ語 | 11,773 | 35,358 |
| 123 | フェロー語 | 10,850 | 26,217 |
| 124 | サハ語 | 10,688 | 32,731 |
| 125 | バイエルン・オーストリア語 | 10,596 | 36,977 |
| 126 | ヴェネト語 | 10,508 | 32,182 |
| 127 | サンスクリット | 10,297 | 30,136 |
| 128 | ナワトル語 | 10,247 | 20,185 |
| 129 | オセット語 | 10,069 | 37,256 |

表 8　Wikipediaの総合計

|  |  |  |
| --- | --- | --- |
| 言語数 | 純記事数 | 総項目数 |
| 288 | 34,213,451 | 128,543,318 |

2014年12月31日現在の試験版のウィキペディアの総項目数上位20の言語

表 9　試験版ウィキペディア

|  |  |  |  |
| --- | --- | --- | --- |
| 順位 | 言語名（日本語） | 言語名（英語） | 総項目数 |
| 1 | 南アゼルバイジャン語 | South Azerbaijani | 3,457 |
| 2 | ロル語 | Luri | 3,196 |
| 3 | カレリア語 | Karelian | 2,098 |
| 4 | ヒリガイノン語 | Hiligaynon | 1,817 |
| 5 | メアンキエリ （トルネダール・フィンランド語） | Meänkieli | 1,812 |
| 6 | 古代ギリシア語 | Ancient Greek | 1,676 |
| 7 | バフティヤーリー語 （ロル語バフティヤーリー方言） | Bakhtiari | 1,636 |
| 8 | タリシュ語 | Talysh | 1,506 |
| 9 | イリガ・ビコール語 | Iriga Bicolano | 1,442 |
| 10 | シルハ語 （タシルハイト語、タシュリヒート語） | Shilha | 1,071 |
| 11 | ユカテコ語 | Yucatec Maya | 994 |
| 12 | 低地シレジア語 （シレジア語ではなくドイツ語の方言） | Silesian German | 989 |
| 13 | アワディー語 | Awadhi | 972 |
| 14 | ゴア・コンカニ語 | Goan Konkani | 901 |
| 15 | リヴォニア語 | Livonian | 747 |
| 16 | 南バローチー語 | Southern Balochi | 731 |
| 17 | ドテリ語 | Doteli | 700 |
| 18 | コワール語 | Khowar | 566 |
| 19 | ラコタ語 | Lakota | 559 |
| 20 | ブラーフーイー語 | Brahui | 544 |

参考文献

[6]Cave cattum (2008) 「Wikipedia:ウィキペディアについて」<http://ja.wikipedia.org/wiki/Wikipedia:%E3%82%A6%E3%82%A3%E3%82%AD%E3%83%9A%E3%83%87%E3%82%A3%E3%82%A2%E3%81%AB%E3%81%A4%E3%81%84%E3%81%A6> (参照2013-7-3)

[7]Sketch (2003) [Help:新規ページの作成」<http://ja.wikipedia.org/wiki/Help:%E6%96%B0%E8%A6%8F%E3%83%9A%E3%83%BC%E3%82%B8%E3%81%AE%E4%BD%9C%E6%88%90> (参照2014-12-16)

[8]G (2003) 「Wikipedia:日本語版の統計」 <http://ja.wikipedia.org/wiki/Wikipedia:%E6%97%A5%E6%9C%AC%E8%AA%9E%E7%89%88%E3%81%AE%E7%B5%B1%E8%A8%88> (参照2015-1-11)

[9]Setu (2003) 「Wikipedia:全言語版の統計」<http://ja.wikipedia.org/wiki/Wikipedia:%E5%85%A8%E8%A8%80%E8%AA%9E%E7%89%88%E3%81%AE%E7%B5%B1%E8%A8%88> (参照2015-1-17)

第4章

調査方法・過程

# 調査・開発



4.1. 本章の構成

本章では，Wikipediaの活動者の編集履歴データの取得の方法，収集したデータより表，グラフの作成過程を述べる．

4.2. 編集回数の多いウィキペディアンのデータ収集について

### 4.2.1. stat.awkで「ファイルの状態」を取得するコマンドを作成する．

BEGIN{s=0;s2=0} {s+=$1;s2+=$1\*$1;} END{m=s/NR;printf("%f,%f\n",m,sqrt(s2/NR-m\*m))}

### 4.2.2. record.awkで「行」のコマンドを作成する．

BEGIN{FS=","}

{

printf("echo %d,$(wget \"http://ja.wikipedia.org/w/index.php?title=特別:投稿記録/%s&limit=500\" -O - | gawk -f diff.awk | gawk -f stat.awk) >> record.csv\n", NR, $2);

}

### pagesize.awkで「ページの大きさ」のコマンドを作成する．

BEGIN{FS=","}

{

printf("echo %d,$(wget \"http://ja.wikipedia.org/w/index.php?title=利用者:%s&action=history\" -O - | gawk -f byte.awk | sed 's/,//g' | sed 's/空/0/g' | head -n 1) >> pagesize.csv\n", NR, $2);

}

### diff.awkで「テキストファイルの差分を抽出する」コマンドを作成する．

/^<li/{print gensub(/.\*\((([+-]([0-9]|,)+)|0)\).\*/, "\\1", $0)}

### byte.awkで「バイトの表示をする」コマンドを作成する．

/^<li>/{print gensub(/.\*\((([0-9,空])+)(バイト)?\).\*/, "\\1", $0)}

## 4.3. Wikipediaの活動調査について

### 4.3.1. 利用者ページのサイズ

「香辛料」というウィキペディアンがいたとする．

この人のページ[http://ja.wikipedia.org/wiki/利用者:香辛料](http://ja.wikipedia.org/wiki/%E5%88%A9%E7%94%A8%E8%80%85:%E9%A6%99%E8%BE%9B%E6%96%99)の履歴表示[http://ja.wikipedia.org/w/index.php?title=利用者:香辛料&action=history](http://ja.wikipedia.org/w/index.php?title=%E5%88%A9%E7%94%A8%E8%80%85:%E9%A6%99%E8%BE%9B%E6%96%99&action=history)を見れば，この人のページのサイズがわかる．

履歴表示をファイルに保存してから確かめる．

wget "http://ja.wikipedia.org/w/index.php?title=利用者:香辛料&action=history" -O tmp

cat tmp\

| gawk -f byte.awk\

| sed 's/,//g'\

| sed 's/空/0/g'\

| head -n 1

ファイルに保存せずに調べるなら，

wget "http://ja.wikipedia.org/w/index.php?title=利用者:香辛料&action=history" -O -\

| gawk -f byte.awk\

| sed 's/,//g'\

| head -n 1

ページがない人(「タイポ女子」のページがないとする)の場合，

wget "http://ja.wikipedia.org/w/index.php?title=利用者:タイポ女子&action=history" -O -\

| gawk -f byte.awk\

| sed 's/,//g'\

| head -n 1

何も出力されない．

### 4.3.2. 投稿記録

「香辛料」の投稿記録[http://ja.wikipedia.org/w/index.php?title=特別:投稿記録/香辛料&limit=500](http://ja.wikipedia.org/w/index.php?title=%E7%89%B9%E5%88%A5:%E6%8A%95%E7%A8%BF%E8%A8%98%E9%8C%B2/%E9%A6%99%E8%BE%9B%E6%96%99&limit=500)を調べてみる．

ファイルに保存して，差分のバイト数だけを抜き出す．

wget "http://ja.wikipedia.org/w/index.php?title=特別:投稿記録/香辛料&limit=500" -O tmp

cat tmp\

| gawk -f diff.awk\

> tmp.csv

cat tmp.csv

平均と標準偏差を求める．

cat tmp\

| gawk -f diff.awk\

| gawk -f stat.awk

ファイルに保存せずに調べるなら，

wget "http://ja.wikipedia.org/w/index.php?title=特別:投稿記録/香辛料&limit=500" -O -\

| gawk -f diff.awk\

| gawk -f stat.awk

### 4.3.3 .編集回数の多いウィキペディアン

[http://ja.wikipedia.org/wiki/Wikipedia:編集回数の多いウィキペディアンの一覧](http://ja.wikipedia.org/wiki/Wikipedia:%E7%B7%A8%E9%9B%86%E5%9B%9E%E6%95%B0%E3%81%AE%E5%A4%9A%E3%81%84%E3%82%A6%E3%82%A3%E3%82%AD%E3%83%9A%E3%83%87%E3%82%A3%E3%82%A2%E3%83%B3%E3%81%AE%E4%B8%80%E8%A6%A7)のテーブルを抜き出す．

python topwikipedians.py > topwikipedians.csv（topwikipedians.csvをExcelで読むと文字化けする．Excelで読みたい場合は，nkf -sjis topwikipedians.csv > topwikipedians.sjis.csvなどとすること．）

Google Spreadsheetで=IMPORTHTML("http://ja.wikipedia.org/wiki/Wikipedia:%E7%B7%A8%E9%9B%86%E5%9B%9E%E6%95%B0%E3%81%AE%E5%A4%9A%E3%81%84%E3%82%A6%E3%82%A3%E3%82%AD%E3%83%9A%E3%83%87%E3%82%A3%E3%82%A2%E3%83%B3%E3%81%AE%E4%B8%80%E8%A6%A7","table",1)としてもよいのだが「(bot)」などを手動で削除しなければならない．

1行目は見出しだから無視し，2行目以降で2列目に入っている名前を抜き出す．

cat topwikipedians.csv\

| tail -n +2\

| gawk -F ',' '{print $2}'

あとは上のスクリプトを再現して実行すればよい．

### 4.3.4. 編集回数の多いウィキペディアンの個人ページサイズ

rm pagesize.csv

cat topwikipedians.csv\

| tail -n +2\

| gawk -f pagesize.awk\

| sh

結果はpagesize.csvに格納される．

### 4.3.5 . 編集回数の多いウィキペディアンの投稿記録

rm record.csv

cat topwikipedians.csv\

| tail -n +2\

| gawk -f record.awk\

| sh

結果はrecord.csvに格納される．

2015年1月22日に取得したデータが20150122\_topwikipedians.csvである．

表 10　Wikipedia：編集回数の多いウィキペディアンランキング

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 順位 | 利用者 | 編集回数 | 総編集回数 | 個人ページサイズ | 編集量の平均 | 編集量の標準偏差 |
| 1 | 編集中 | 2699 | 12863 | 777 | 0.058116 | 111.028922 |
| 2 | Kamaura | 1758 | 2931 | 24 | 231.762 | 170.229073 |
| 3 | タイポ女子 | 1739 | 4506 | 0 | 33.984 | 43.91915 |
| 4 | そらみみ | 1556 | 5195 | 891 | 54.635815 | 119.865108 |
| 5 | Eurotuber | 1545 | 15893 | 8793 | 63.326 | 125.415309 |
| 6 | まっきんりい | 1361 | 3013 | 0 | -12.486974 | 116.610554 |
| 7 | 60.62.156.41 | 1300 | 1300 | 0 | 9.266 | 137.997273 |
| 8 | Ihimutefu | 1262 | 49837 | 0 | 12.132 | 44.592853 |
| 9 | 221.94.14.18 | 1199 | 26678 | 0 | 56.666 | 96.732551 |
| 10 | TOYO GTO | 1058 | 12661 | 465 | 22.534 | 195.91334 |
| 11 | 曾禰越後守 | 1034 | 56763 | 10543 | 37.594 | 84.135469 |
| 12 | Banku | 1013 | 7869 | 1078 | 40.048387 | 116.255774 |
| 13 | 119.171.132.79 | 951 | 11995 | 0 | 56.693548 | 243.841639 |
| 14 | ふーくん | 922 | 3618 | 0 | 10.334669 | 174.63108 |
| 15 | Dr.KRX | 916 | 9717 | 139 | 1.97 | 85.860148 |
| 16 | Yamato-i | 912 | 40236 | 48 | 33.560241 | 111.050937 |
| 17 | Sidiouschancellor | 898 | 898 | 1614 | -30.798 | 109.246113 |
| 18 | Garakmiu2 | 851 | 8079 | 12167 | 124.651303 | 205.409427 |
| 19 | MiraModreno | 826 | 47740 | 41 | 81.566 | 149.050507 |
| 20 | Championship2006 | 804 | 104706 | 0 | 74.895582 | 125.642706 |
| 21 | Takatakabouzu | 798 | 1964 | 0 | 11.49 | 118.292053 |
| 22 | Norio-abc | 773 | 10804 | 1038 | 30.442 | 64.101097 |
| 23 | Stealth t | 763 | 4682 | 3235 | 91.223849 | 163.886499 |
| 24 | Dokudami | 712 | 1170 | 249 | 73.472946 | 124.932642 |
|  | Love monju bot | 704 | 1365 | 1396 | 7.454 | 30.194302 |
| 25 | 101.142.164.224 | 702 | 702 | 0 | 133.372 | 194.131939 |
| 26 | Zazanasawa | 698 | 9337 | 653 | 53.767535 | 97.66446 |
| 27 | Rain night-AWB | 694 | 14113 | 1155 | 11.22 | 23.422289 |
| 28 | Vielen dank | 686 | 790 | 559 | 92.47 | 166.668801 |
| 29 | るなあるてみす | 684 | 75419 | 2544 | 22.366 | 69.471117 |
| 29 | Otohakusyu | 684 | 6168 | 609 | 41.984 | 83.292927 |
| 31 | Kekero | 675 | 2941 | 1089 | 38.148 | 86.780217 |
| 32 | Cueeibu | 647 | 839 | 235 | -6.082164 | 148.027983 |
| 33 | Frozen-mikan | 643 | 9891 | 3381 | 18.382 | 207.922053 |
| 34 | YOUJI | 634 | 12910 | 21776 | 92.196 | 149.671739 |
| 35 | 113.38.179.234 | 631 | 1494 | 0 | 13.232 | 105.346486 |
| 36 | 神奈川エーフレッツ | 612 | 2809 | 0 | 23.661458 | 232.021169 |
| 37 | Tokoma191919 | 607 | 7896 | 0 | 53.228 | 99.21639 |
| 38 | レタス | 598 | 39453 | 0 | 126.848 | 61.739298 |
| 39 | Umitori | 596 | 8818 | 0 | 108.558468 | 227.911221 |
| 40 | Physalis-02A | 590 | 6736 | 1506 | 97.150301 | 199.118326 |
| 41 | 桂鷺淵 | 577 | 19220 | 3124 | 76.346 | 172.324189 |
| 42 | Tiyoringo | 572 | 77190 | 6224 | 88.610442 | 255.459797 |
| 43 | 香辛料 | 550 | 26548 | 2570 | 148.370741 | 196.249639 |
| 44 | 210a | 549 | 11493 | 480 | 23.723447 | 145.284407 |
| 45 | Oorong-tea | 539 | 5677 | 0 | 134.720648 | 282.461475 |
| 46 | 2407:C800:34F0:A:0:0:2:F0BF | 538 | 538 | 0 | 76.108 | 145.805351 |
| 47 | たいすけ55 | 534 | 8189 | 14073 | 49.173387 | 151.295741 |
| 48 | Tahuantinsuyo | 530 | 568 | 547 | 184.715431 | 212.158671 |
| 49 | Ef3 | 527 | 4118 | 26885 | 142.004 | 237.726322 |
| 50 | Nissk | 524 | 16155 | 0 | 85.556 | 166.907264 |
| 51 | NOBU | 517 | 5452 | 658 | 151.564257 | 263.53493 |
| 52 | More blessing | 515 | 14972 | 28463 | 117.8 | 185.06481 |
| 53 | XRGD | 496 | 10925 | 0 | 69.219316 | 199.849416 |
| 54 | Johnny Gan | 495 | 523 | 0 | 113.054 | 88.845344 |
| 55 | Yanajin33 | 491 | 7852 | 5392 | 46.402 | 43.783335 |
| 56 | Dr jimmy | 489 | 124648 | 251 | -57.384615 | 257.831964 |
| 57 | Arvin | 484 | 2656 | 1725 | 47.383534 | 164.591802 |
| 58 | スカラマンガ | 478 | 10019 | 591 | 51.364 | 157.313774 |
| 59 | JungleCrow | 469 | 5324 | 5478 | 54.406814 | 235.274616 |
| 60 | CommonsDelinker | 468 | 39729 | 7681 | -45.018 | 60.731159 |
| 61 | Terumasa | 466 | 8806 | 0 | 131.232932 | 193.971176 |
| 62 | 180.16.155.30 | 455 | 1942 | 0 | 50.448 | 127.610733 |
| 63 | チンドレ・マンドレ | 447 | 13207 | 4854 | 174.382294 | 297.011288 |
| 64 | 沙馬琉寛土 | 444 | 1763 | 377 | 132.726 | 158.542067 |
| 65 | Pekanpe | 441 | 2555 | 5988 | 83.957746 | 201.063511 |
| 66 | Puronohu | 435 | 761 | 30 | 75.247485 | 134.09976 |
| 67 | H ichi34 | 425 | 2579 | 0 | 64.396 | 58.723685 |
| 68 | タケナカ | 416 | 20107 | 684 | 181.382231 | 339.268252 |
| 69 | 山川草木 | 413 | 3913 | 0 | 37.992 | 146.3435 |
| 70 | Enyokoyama | 407 | 4798 | 0 | 12.263485 | 143.552306 |
| 71 | 極楽サタン | 400 | 15053 | 0 | 22.786 | 64.869409 |
| 72 | TempuraDON | 396 | 6674 | 69 | 70.772 | 174.055244 |
| 73 | JapaneseA | 395 | 5161 | 3369 | 177.702213 | 304.280462 |
| 74 | けぶお | 392 | 2851 | 0 | 62.969697 | 113.16155 |
| 75 | ミラー海神 | 389 | 3047 | 0 | 71.53629 | 92.750634 |
| 76 | 花蝶風月雪月花警部 | 386 | 38396 | 12 | 13.322 | 218.535906 |
| 77 | HOPE | 382 | 19809 | 2892 | 85.305835 | 256.500076 |
| 78 | 須磨寺横行 | 380 | 11882 | 498 | -7.963928 | 316.527272 |
| 79 | Suikotei | 377 | 7517 | 0 | 49.792 | 172.189642 |
| 80 | しっち | 369 | 2993 | 0 | 86.266 | 105.703771 |
| 80 | 210.162.60.12 | 369 | 2868 | 0 | -4.922 | 22.77033 |
| 82 | HTTH | 367 | 1412 | 4798 | 63.438776 | 152.609216 |
| 83 | とほほのほ | 364 | 2648 | 158 | 116.768 | 198.800418 |
| 84 | Atsonivt | 360 | 6401 | 0 | 40.745455 | 120.810294 |
| 85 | みそ煮込みうどん | 357 | 1126 | 0 | 63.77666 | 121.162397 |
| 86 | 頭蓋骨抉る | 351 | 4170 | 0 | 212.516 | 207.587114 |
| 87 | Hruygo | 347 | 16985 | 1120 | 123.17 | 255.077708 |
| 87 | 山の山手線 | 347 | 5167 | 5743 | 81.926 | 122.183045 |
| 89 | Kurirai1 | 345 | 4309 | 0 | 108.58 | 183.021506 |
| 90 | 松茸 | 340 | 58631 | 8973 | 132.40481 | 200.528541 |
| 91 | LearningBox | 339 | 18796 | 2906 | 77.532129 | 218.532813 |
| 92 | モンゴルの白い虎 | 335 | 31560 | 0 | 28.016 | 158.690402 |
| 92 | Unamu | 335 | 27196 | 998 | 86.16 | 192.293844 |
| 92 | Ryun | 335 | 815 | 28 | -49.357143 | 178.662334 |
| 95 | サカガミタロウ | 332 | 9641 | 897 | 187.839679 | 261.192829 |
| 96 | Sumaru | 331 | 6622 | 8855 | 113.082 | 190.642218 |
| 97 | ケイ20003 | 330 | 7678 | 0 | 79.356 | 184.545575 |
| 97 | アストロニクル | 330 | 2816 | 708 | 126.894 | 225.11133 |
| 99 | 2407:C800:34F0:A:0:0:2:FC22 | 329 | 329 | 0 | 29.736364 | 96.665247 |
| 100 | Softyu | 326 | 4284 | 0 | 69.695565 | 225.876702 |
| 101 | Rojet | 324 | 1228 | 0 | 135.568273 | 190.778569 |
| 102 | JR1NVB | 323 | 10350 | 2585 | 48.454 | 90.091686 |
| 103 | 220.247.111.83 | 322 | 398 | 0 | 139.220441 | 180.176699 |
| 104 | Kashima51 | 316 | 4488 | 2354 | 149.768 | 190.217671 |
| 105 | Tonbi ko | 315 | 1294 | 17055 | 128.425703 | 265.942296 |
| 105 | 153.192.100.232 | 315 | 615 | 0 | 0.388 | 49.055616 |
| 105 | Gagagagazette26 | 315 | 589 | 53 | 53.983903 | 156.716663 |
| 108 | 中竹たけし | 314 | 3132 | 2137 | 99.344758 | 181.541795 |
| 109 | Gowithitjam | 309 | 309 | 131 | 89.253036 | 291.874049 |
| 110 | MaximusM4 | 303 | 12053 | 3505 | 18.55511 | 159.396817 |
| 111 | いいすく | 302 | 26195 | 707 | 26.362 | 160.733441 |
| 111 | 58.0.75.79 | 302 | 303 | 0 | 41.605991 | 128.446654 |
| 113 | 220.147.165.60 | 300 | 386 | 0 | 37.280561 | 84.279627 |
| 114 | Hyuki-h | 298 | 9464 | 0 | 52.912 | 107.260339 |
| 114 | 134.208.33.109 | 298 | 686 | 0 | 88.468 | 153.301406 |
| 116 | KASEI | 295 | 14728 | 3706 | 78.9 | 188.001367 |
| 116 | タバコはマーダー | 295 | 1897 | 0 | 82.240891 | 235.709259 |
| 116 | 123.230.197.118 | 295 | 295 | 0 | 175.04 | 228.308621 |
| 119 | 馳夫 | 293 | 293 | 1337 | -21.82 | 85.875699 |
|  | Tribot | 290 | 212817 | 20830 | 25.65 | 23.360383 |
| 120 | SSC | 290 | 21451 | 170 | 88.968 | 163.015518 |
| 121 | DEBUZ | 281 | 791 | 0 | 139.010331 | 247.548071 |
| 121 | Corpsebot | 281 | 399 | 594 | 4.988 | 23.157631 |
| 123 | Fielder | 278 | 5974 | 410 | 61.752 | 262.432846 |
| 124 | Inglid | 275 | 300 | 78 | 154.018256 | 235.161102 |
| 125 | Ktp625 | 273 | 4440 | 0 | 117.8107 | 263.545947 |
| 126 | Dɐ | 270 | 2139 | 2694 | 94.73306 | 163.633331 |
| 127 | Geogie | 268 | 37535 | 3178 | 117.29697 | 200.599371 |
| 128 | 美区 | 267 | 3961 | 70 | 106.341053 | 213.463493 |
| 128 | 布哇 | 267 | 2852 | 967 | 17.174699 | 283.357501 |
| 130 | 兼六園ランド | 266 | 425 | 0 | 133.822102 | 236.144756 |
| 131 | カレー風味ハヤシライス | 265 | 19639 | 13307 | 80.15 | 125.733828 |
| 131 | Kekusnek | 265 | 3470 | 0 | 83.631048 | 134.77774 |
| 133 | 超プロ住民 | 264 | 32225 | 0 | 61.945455 | 125.56532 |
| 134 | 山田晴通 | 263 | 16205 | 77325 | 86.14 | 174.081017 |
| 135 | Aoiya | 258 | 25050 | 87 | 67.732252 | 189.80101 |
| 136 | 2407:C800:34F0:A:0:0:2:EC0B | 257 | 257 | 0 | 10.916031 | 93.065932 |
| 137 | Xapones | 256 | 15209 | 300 | 49.200811 | 173.097524 |
| 137 | Benichan | 256 | 4325 | 0 | 209.843373 | 243.560132 |
| 137 | ワーナー成増 | 256 | 1373 | 78415 | 52.128 | 216.488345 |
| 140 | JamvBot | 255 | 263 | 664 | 49.1 | 111.838811 |
| 141 | 122.17.103.226 | 252 | 848 | 0 | 85.068 | 163.540378 |
| 142 | Bonnie lepps | 246 | 3087 | 0 | 167.412475 | 223.131808 |
| 142 | 60.37.119.126 | 246 | 970 | 0 | 196.285141 | 228.030116 |
| 144 | 欣之介くん | 244 | 397 | 59 | 66.206897 | 216.180487 |
| 145 | さかおり | 243 | 13498 | 20234 | 29.286 | 253.894128 |
| 146 | Scoutsscouts | 240 | 443 | 0 | 14.954955 | 51.809489 |
|  | BankuBot | 239 | 43996 | 70 | 8.208 | 26.460173 |
| 147 | M9106TB | 239 | 10757 | 0 | 86.943888 | 217.651983 |
| 148 | 121.116.225.167 | 238 | 238 | 0 | 59.066 | 132.183061 |
| 149 | 219.43.134.94 | 237 | 12332 | 0 | 57.23 | 93.340779 |
| 149 | Ackeyyama | 237 | 6628 | 0 | 64.080321 | 158.569729 |
| 151 | Banpei | 236 | 1499 | 0 | 90.384 | 149.840384 |
| 151 | Homura magica | 236 | 542 | 58 | 155.642857 | 239.215386 |
| 153 | 上州力 | 234 | 9509 | 1043 | 33.922 | 51.327224 |
| 154 | Lovesouleyes | 232 | 8846 | 1336 | 101.30101 | 262.077204 |
| 155 | コシャリ | 228 | 8086 | 0 | 156.78 | 219.077711 |
| 155 | OskNe | 228 | 3751 | 553 | 49.16 | 220.827268 |
| 157 | 博多商人コイツミ | 222 | 15882 | 6988 | 182.051335 | 307.586344 |
| 158 | Tomato tomato tomato | 220 | 224 | 1030 | -131.126761 | 235.656748 |
| 159 | Harima49 | 219 | 5724 | 0 | -4.9 | 30.380026 |
| 159 | ノフノフ | 219 | 2440 | 23288 | 48.757576 | 152.199158 |
| 159 | 180.2.164.59 | 219 | 219 | 0 | 191.568376 | 241.112197 |
| 162 | 新家 | 215 | 1041 | 0 | 43.642 | 85.031723 |
| 163 | TAMA | 213 | 15078 | 0 | 6.466 | 85.746678 |
| 163 | Shinsh | 213 | 1631 | 196 | 143.566 | 201.075274 |
| 165 | Miraculous | 210 | 210 | 55 | 37.134868 | 24.506945 |
| 166 | Quark Logo | 208 | 1647 | 3460 | 117.29878 | 270.955382 |
| 166 | ☆あらは | 208 | 1404 | 2506 | 31.747967 | 165.886549 |
| 168 | ゼロツー | 207 | 15628 | 709 | 43.102 | 172.469451 |
| 168 | 豆もち | 207 | 2805 | 0 | -16.274 | 136.559683 |
| 168 | ShuBraque | 207 | 208 | 1808 | 62.215078 | 182.152869 |
| 168 | Meneaki1283 | 207 | 207 | 0 | 54.438356 | 169.203964 |
| 172 | Iceandsnow7 | 205 | 2278 | 0 | 101.627291 | 223.012484 |
| 172 | 孫子兵法 | 205 | 294 | 0 | 81.389522 | 230.329072 |
| 174 | Platinum-phantom | 202 | 2162 | 30838 | 220.418367 | 298.715041 |
| 175 | Chiba ryo | 201 | 11586 | 91671 | 123.995976 | 225.895268 |
| 175 | WEWEE | 201 | 7391 | 0 | 29.53507 | 125.880993 |
| 175 | Assemblykinematics | 201 | 372 | 26865 | 176.597561 | 296.201372 |
| 178 | Komaru | 200 | 13826 | 0 | 209.247465 | 246.540847 |
| 178 | 124.33.203.210 | 200 | 825 | 0 | 139.266533 | 217.887579 |
| 178 | 221.118.87.67 | 200 | 378 | 0 | 64.574 | 154.795144 |
| 181 | Magyon | 198 | 18758 | 3654 | 108.417671 | 245.717592 |
| 181 | Zezezozo | 198 | 409 | 0 | 51.804 | 75.095443 |
| 181 | Kbgks | 198 | 269 | 0 | 103.612 | 196.384097 |
| 181 | Namatamago195 | 198 | 242 | 454 | 37.417219 | 97.462668 |
| 181 | 219.97.22.191 | 198 | 198 | 0 | 113.703518 | 137.325486 |
| 186 | Muyo | 197 | 30028 | 57171 | 56.756539 | 290.838922 |
| 186 | 麒麟檸檬 | 197 | 2520 | 0 | 33.03006 | 192.936046 |
| 186 | ホージー | 197 | 433 | 0 | 129.785714 | 245.306275 |
| 189 | Wjp28y | 196 | 9897 | 17651 | 57.110887 | 196.606948 |
| 189 | Menetoko | 196 | 880 | 1507 | 90.122244 | 256.447847 |
| 191 | ＫｏＺ | 194 | 10061 | 1249 | 177.818548 | 284.579598 |
| 191 | Kanade tachibana AB! | 194 | 258 | 0 | 88.109551 | 178.733581 |
| 193 | Lanwi1 | 193 | 1583 | 904 | 61.596386 | 285.559649 |
| 194 | SNOW RIDE | 190 | 460 | 0 | 60.102204 | 120.717244 |
| 195 | タペストリー | 189 | 1351 | 0 | 20.193548 | 223.568788 |
| 195 | 202.35.124.66 | 189 | 963 | 0 | -13.276 | 48.052886 |
| 197 | 蓮音 | 188 | 56058 | 42 | 185.548 | 121.465055 |
| 198 |  | 187 | 4748 | 2327 | 139.300203 | 286.260097 |
| 199 | Nova replet laetitia | 186 | 186 | 20 | 77.813559 | 195.539365 |
| 200 | Gtorew | 184 | 2794 | 0 | 65.307847 | 175.463651 |

(期間2014-07-01 – 2014-07-31 )

### 4.3.6. 編集回数の多いウィキペディアンのヒストグラム

Excelでヒストグラムの作り方について述べる．

各項目での範囲を決めるために最大値と最小値を出す．

表 11　最大値と最小値の表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **編集回数** | **総編集回数** | **個人ページサイズ** | **編集量の平均** | **編集量の標準偏差** |
| 最大値 | 2699 | 212817 | 91671 | 231.762 | 339.268252 |
| 最小値 | 184 | 186 | 0 | -131.126761 | 22.77033 |

Excelを開いてから，「ファイル」→「オプション」→「アドイン」→「設定」から「分析ツール」にチェックを付けてから，「OK」をクリックする．「データ」のタブに「分析」が加わる．「分析」の中に「分析データ」が入る．赤い枠で示す．

Excelを開いてから，「ファイル」をクリックし，その中にある「オプション」をクリックする．



図 2　ファイルからオプション

「オプション」が開くと中に「アドイン」があるのでクリックする．

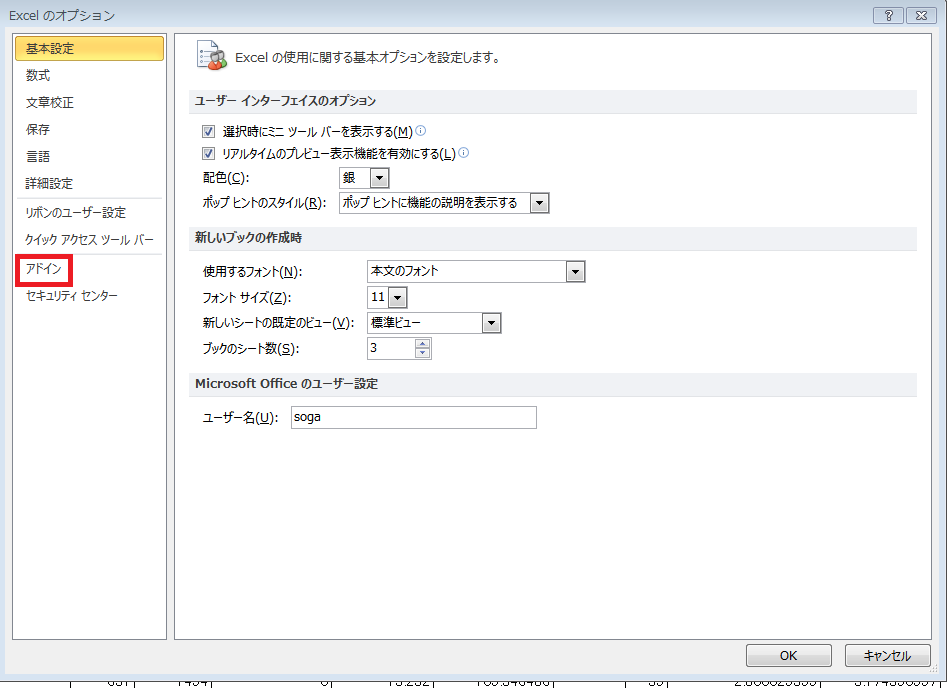


図 3　オプションからアドイン

「アドイン」を開くと下部に「設定」が表示されるのでクリックする．

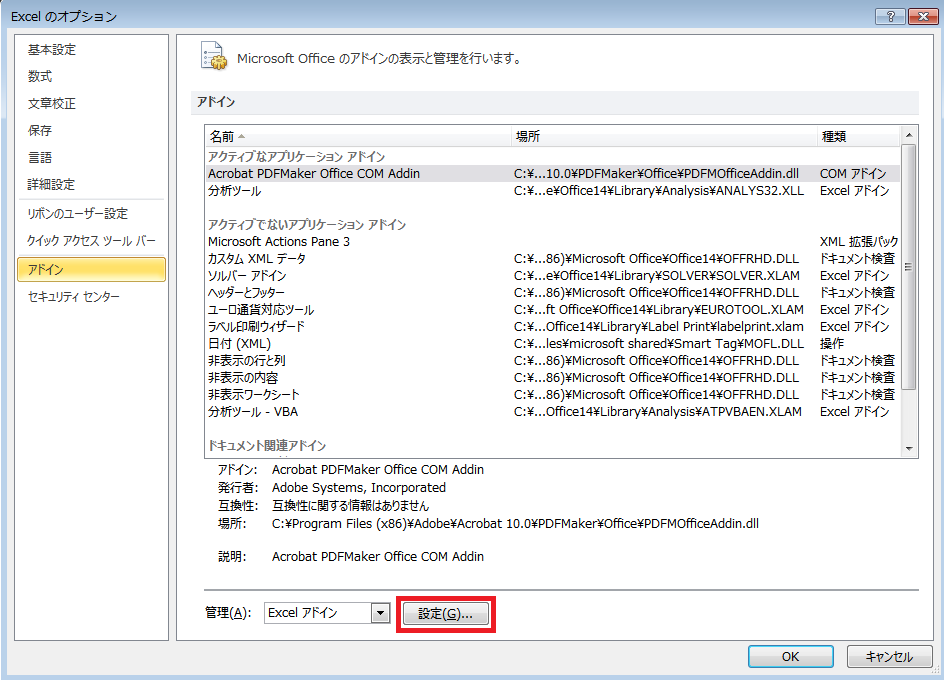


図 4　アドインから設定

「設定」を開いて，「分析ツール」をクリックしてチェックをいれる．

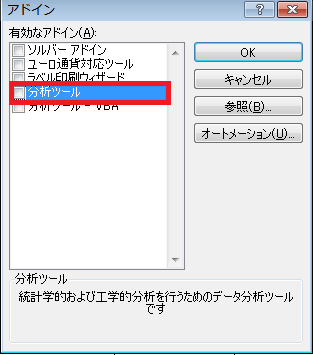


図 5　設定から分析ツール

チェックを入れたら「OK」をクリックする．そうするとExcelの「データ」タブに「分析」の欄が入る．「分析」の欄に「データ分析」が入っているのでクリックする．すると以下の図のようにダイアログボックスが表示されるので，「ヒストグラム」を選択し「OK」をクリックする．

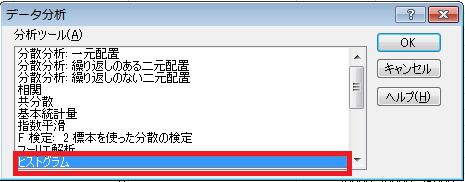


図 6　データ分析からヒストグラムを選択

ヒストグラムのダイアログボックスが表示されたら，「入力範囲」に求めたいデータの数値の列を入力し，「データ区間」に求めたい階級の上限の数値の列を入力する．ヒストグラムのグラフを作成したいので「グラフ作成」にチェックをいれ，「OK」をクリックする．

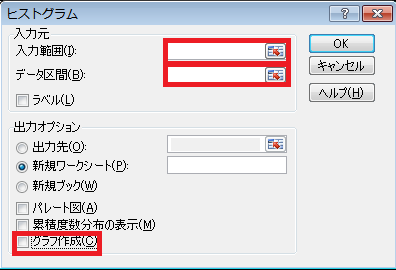


図 7　入力範囲，データ区間に入力

ヒストグラムのグラフと「入力範囲」と「データ区間」に入力した表がExcelに表示される．

ヒストグラムのグラフが以下の図では不完全なため，グラフの部分にマウスを持っていき，左クリックをする．

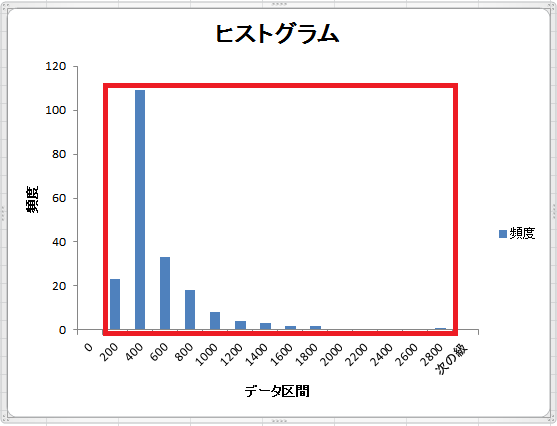


図 8　不完全のヒストグラム

左クリックをすると「データ系列の書式設定」をクリックする．

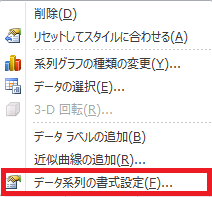


図 9 データ系列の書式設定

系列のオプションから，「要素の間隔」を「なし」に変更し閉じる．

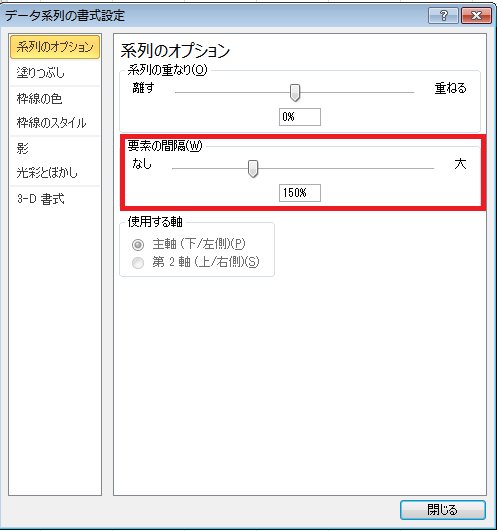


図 10　データ系列の書式設定

要素の間隔を「なし」にするとヒストグラムになる．

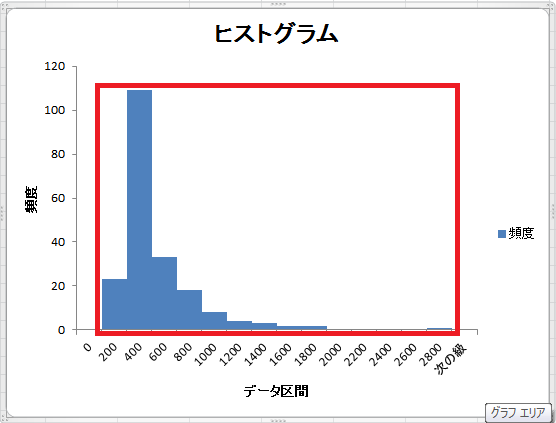


図 11　ヒストグラム

Excelの「データ分析」のヒストグラムだと整数値に限ればいいが，実数値にはむいていないので，COUNTIFSを使って求めることにした．

Excelで範囲を=COUNTIFSを用いて条件に合うセルを指定すると，以下の図のようになる．

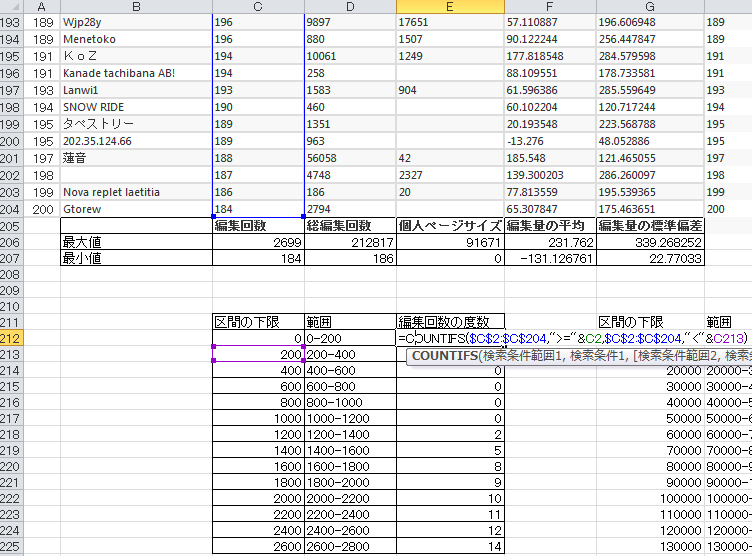


図 12　COUNTIFSで範囲指定

次に区切る範囲を決める．

図12の様にそれぞれ範囲を決め度数表を作成する．また数値が大きいので，常用対数の「log10」を用いた度数表も作成する．

表 12　編集回数の度数表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | 編集回数の度数 |
| 0 | 0-200 | 20 |
| 200 | 200-400 | 111 |
| 400 | 400-600 | 34 |
| 600 | 600-800 | 18 |
| 800 | 800-1000 | 8 |
| 1000 | 1000-1200 | 4 |
| 1200 | 1200-1400 | 3 |
| 1400 | 1400-1600 | 2 |
| 1600 | 1600-1800 | 2 |
| 1800 | 1800-2000 | 0 |
| 2000 | 2000-2200 | 0 |
| 2200 | 2200-2400 | 0 |
| 2400 | 2400-2600 | 0 |
| 2600 | 2600-2800 | 1 |
| 2800 | 2800- | 0 |

表 13　log10(編集回数の度数)表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | log10(編集回数) |
| 0 | 0-0.5 | 0 |
| 0.5 | 0.5-1 | 0 |
| 1 | 1-1.5 | 0 |
| 1.5 | 1.5-2 | 0 |
| 2 | 2-2.5 | 99 |
| 2.5 | 2.5-3 | 92 |
| 3 | 3-3.5 | 12 |
| 3.5 | 3.5-4 | 0 |
| 4 | 4-4.5 | 0 |
| 4.5 | 4.5-5 | 0 |
| 5 | 5-5.5 | 0 |
| 5.5 | 5.5- | 0 |

表 14　総編集回数の度数表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | 総編集回数の度数 |
| 0 | 0-10000 | 142 |
| 10000 | 10000-20000 | 35 |
| 20000 | 20000-30000 | 7 |
| 30000 | 30000-40000 | 7 |
| 40000 | 40000-50000 | 4 |
| 50000 | 50000-60000 | 3 |
| 60000 | 60000-70000 | 0 |
| 70000 | 70000-80000 | 2 |
| 80000 | 80000-90000 | 0 |
| 90000 | 90000-100000 | 0 |
| 100000 | 100000-110000 | 1 |
| 110000 | 110000-120000 | 0 |
| 120000 | 120000-130000 | 1 |
| 130000 | 130000-140000 | 0 |

～～～～～～～～～～～～～～～～～～～～～

|  |  |  |
| --- | --- | --- |
| 190000 | 190000-200000 | 0 |
| 200000 | 200000-210000 | 0 |
| 210000 | 210000-220000 | 1 |
| 220000 | 220000- | 0 |

表 15 log10(総編集回数の度数)表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | log10(総編集回数) |
| 0 | 0-0.5 | 0 |
| 0.5 | 0.5-1 | 0 |
| 1 | 1-1.5 | 0 |
| 1.5 | 1.5-2 | 0 |
| 2 | 2-2.5 | 19 |
| 2.5 | 2.5-3 | 31 |
| 3 | 3-3.5 | 43 |
| 3.5 | 3.5-4 | 49 |
| 4 | 4-4.5 | 44 |
| 4.5 | 4.5-5 | 14 |
| 5 | 5-5.5 | 3 |
| 5.5 | 5.5- | 0 |

表 16 個人ページサイズの度数表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | 個人ページサイズの度数 |
| 0 | 0-5000 | 91 |
| 5000 | 5000-10000 | 10 |
| 10000 | 10000-15000 | 4 |
| 15000 | 15000-20000 | 2 |
| 20000 | 20000-25000 | 4 |
| 25000 | 25000-30000 | 3 |
| 30000 | 30000-35000 | 1 |
| 35000 | 35000-40000 | 0 |
| 40000 | 40000-45000 | 0 |
| 45000 | 45000-50000 | 0 |
| 50000 | 50000-55000 | 0 |
| 55000 | 55000-60000 | 1 |
| 60000 | 60000-65000 | 0 |
| 65000 | 65000-70000 | 0 |
| 70000 | 70000-75000 | 0 |
| 75000 | 75000-80000 | 2 |
| 80000 | 80000-85000 | 0 |
| 85000 | 85000-90000 | 0 |
| 90000 | 90000-95000 | 1 |
| 95000 | 95000- | 0 |

表 17 log10(個人ページサイズの度数)表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | log10(個人ページサイズ) |
| 0 | 0-0.5 | 87 |
| 0.5 | 0.5-1 | 0 |
| 1 | 1-1.5 | 5 |
| 1.5 | 1.5-2 | 12 |
| 2 | 2-2.5 | 9 |
| 2.5 | 2.5-3 | 25 |
| 3 | 3-3.5 | 27 |
| 3.5 | 3.5-4 | 20 |
| 4 | 4-4.5 | 14 |
| 4.5 | 4.5-5 | 4 |
| 5 | 5-5.5 | 0 |
| 5.5 | 5.5- | 0 |

表 18 編集量の度数表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | 編集量の平均の度数 |
| -150 | -150—100 | 1 |
| -100 | -100--50 | 1 |
| -50 | -50-0 | 11 |
| 0 | 0-50 | 59 |
| 50 | 50-100 | 72 |
| 100 | 100-150 | 37 |
| 150 | 150-200 | 17 |
| 200 | 200-250 | 5 |
| 250 | 250- | 0 |

表 19 log10(編集量の度数)表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | log10(編集量の平均) |
| -1.5 | -1.5--1 | 1 |
| -1 | -1--0.5 | 0 |
| -0.5 | -0.5-0 | 1 |
| 0 | 0-0.5 | 14 |
| 0.5 | 0.5-1 | 5 |
| 1 | 1-1.5 | 25 |
| 1.5 | 1.5-2 | 98 |
| 2 | 2-2.5 | 59 |
| 2.5 | 2.5-3 | 0 |
| 3 | 3-3.5 | 0 |
| 3.5 | 3.5-4 | 0 |
| 4 | 4-4.5 | 0 |
| 4.5 | 4.5-5 | 0 |
| 5 | 5-5.5 | 0 |
| 5.5 | 5.5- | 0 |

表 20 編集量の標準偏差の度数表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | 編集量の標準偏差の度数 |
| 0 | 0-50 | 13 |
| 50 | 50-100 | 27 |
| 100 | 100-150 | 39 |
| 150 | 150-200 | 55 |
| 200 | 200-250 | 42 |
| 250 | 250-300 | 23 |
| 300 | 300-350 | 4 |
| 350 | 350- | 0 |

表 21 log10(編集量の標準偏差の度数)表

|  |  |  |
| --- | --- | --- |
| 区間の下限 | 範囲 | log10(編集量の標準偏差) |
| 0 | 0-0.5 | 0 |
| 0.5 | 0.5-1 | 0 |
| 1 | 1-1.5 | 8 |
| 1.5 | 1.5-2 | 32 |
| 2 | 2-2.5 | 161 |
| 2.5 | 2.5-3 | 2 |
| 3 | 3-3.5 | 0 |
| 3.5 | 3.5-4 | 0 |
| 4 | 4-4.5 | 0 |
| 4.5 | 4.5-5 | 0 |
| 5 | 5-5.5 | 0 |
| 5.5 | 5.5- | 0 |

参考文献

[9]Ｇ. (2003) 「Wikipedia:日本語版の統計」 <http://ja.wikipedia.org/wiki/Wikipedia:%E6%97%A5%E6%9C%AC%E8%AA%9E%E7%89%88%E3%81%AE%E7%B5%B1%E8%A8%88>　(参照2014-10-11)

[10]奥村　晴彦 (2014) 「Excelでヒストグラム(度数分布図)」　<http://oku.edu.mie-u.ac.jp/~okumura/stat/excel-hist.html>　(2014-07-29)

[11]秋雪. (2006)　「Wikipedia:編集回数の多いウィキペディアンの一覧」 <http://ja.wikipedia.org/w/index.php?title=Wikipedia:%E7%B7%A8%E9%9B%86%E5%9B%9E%E6%95%B0%E3%81%AE%E5%A4%9A%E3%81%84%E3%82%A6%E3%82%A3%E3%82%AD%E3%83%9A%E3%83%87%E3%82%A3%E3%82%A2%E3%83%B3%E3%81%AE%E4%B8%80%E8%A6%A7&action=info>　(参照2014-8-11)

第5章

調査結果・考察

# 調査結果考察



## 5.1. 本章の構成

　本章では，前章でのWikipedia調査の結果を記述し，その結果に対する考察を記述する．

## 5.2. ヒストグラムの結果

前章で作成したヒストグラムの度数表を基に，ヒストグラムのグラフを作成し，それに対しての考察を述べる．

図 13 編集回数の度数ヒストグラム

編集回数の度数ヒストグラムからランキングに入っている200人の半数以上の人は1ヶ月の間に200～400回程度の編集を行っていることがわかる．編集回数が400回を超えると編集者の数が減ってしまっている．

図 14 総編集回数の度数ヒスとグラム

総編集回数の度数では，10000回未満の人が約七割程度となっており，約2割の利用者しか20000回を超える編集を行っていないのがわかる．

図 15 個人ページサイズの度数ヒストグラフ

個人ページサイズの度数では，約4割の人はページが削除されてしまっていたり，ページを作成していない人たちであったため0kbとなっていた．個人ページを編集している人の約5割は5000kb以下のページであることがわかる．

図 16　編集量の平均の度数ヒストグラム

編集量の平均の度数では，約9割の人たちが記事を編集してページを大きくしていることがわかる．約3割の人たちは50～100ほどページを大きくしている人たちである．一部の人はページを編集して短くしていることを図から読み取ることができる．

図 17　編集量の標準偏差度数ヒストグラム

　編集量の標準偏差の平均が168.72609であったので，約3割の人は200人の中でも編集量が多いという事がわかる．編集量の標準偏差が300を超える人はごく1部でWikipediaによく貢献している．

前章で得たデータを常用対数のlog10を使って表した表である.

表 22　log10の度数表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 下限の区間 | 範囲 | **log10(編集回数)** | **log10(総編集回数)** | **log10(個人ページサイズ)** | **log10(編集量の平均)** | **log10(編集量の標準偏差)** |
| -1.5 | -1.5--1 |  |  |  | 1 |  |
| -1 | -1--0.5 |  |  |  | 0 |  |
| -0.5 | -0.5-0 |  |  |  | 1 |  |
| 0 | 0-0.5 | 0 | 0 | 87 | 14 | 0 |
| 0.5 | 0.5-1 | 0 | 0 | 0 | 5 | 0 |
| 1 | 1-1.5 | 0 | 0 | 5 | 25 | 8 |
| 1.5 | 1.5-2 | 0 | 0 | 12 | 98 | 32 |
| 2 | 2-2.5 | 99 | 19 | 9 | 59 | 161 |
| 2.5 | 2.5-3 | 92 | 31 | 25 | 0 | 2 |
| 3 | 3-3.5 | 12 | 43 | 27 | 0 | 0 |
| 3.5 | 3.5-4 | 0 | 49 | 20 | 0 | 0 |
| 4 | 4-4.5 | 0 | 44 | 14 | 0 | 0 |
| 4.5 | 4.5-5 | 0 | 14 | 4 | 0 | 0 |
| 5 | 5-5.5 | 0 | 3 | 0 | 0 | 0 |
| 5.5 | 5.5- | 0 | 0 | 0 | 0 | 0 |

この表を見ると数値の大体が区間の「2」の周辺であることがわかる.log10(編集回数)とlog10(総編集回数)が0~2までの間が0に対して，log10(編集量の平均)とlog10(編集量の標準偏差)は3~5.5-までが0になっていて逆の変化をしているように見える．

Log10にしたヒストグラムを以下に示す．

図 18　log10(編集回数)の度数ヒストグラム

Log10で編集回数の頻度を撮ってみると約9割の人が2-3の間に居てlog10にする前のヒストグラムと比べると0-800回程度の編集をしていた人たちだとわかる、．

図 19　log10(総編集回数)の度数ヒストグラム

総編集回数が0-10000回までの人たちが大半を占めていたが，log10にしてみてみると，大体の人たちの総編集回数に大きな差はそれほどないようにうかがえる．

図 20　log10(個人ページサイズ)の度数ヒストグラム

個人ページサイズのlog10にする前では0-50000kbの人が大半を占めていて5000kb以上の人はそんなに多くないという結果だった．log10にしてみると，やはり大半を占めているのは帆人ページサイズの大きくない人たちであるが，個人ページサイズの多い人も居ることがわかる.

図 21 log10(編集量の平均)の度数ヒストグラム

編集量の平均でもみられたページサイズを少なくしている人もいるが，約9割以上の人はページサイズが大きくなるように編集した量が多い人がほとんどである.

図 22　log10(編集量の標準偏差)の度数ヒストグラム

編集量の標準偏差ではきれいな山なりになるようなヒストグラムになっていたがlog10にしてみるとバランスが悪く偏っっていることが読み取る事ができる.

図 23　log10の各度数の折れ線グラフ

このグラフはlog10にした200人のデータを一つにまとめたものである.演習量の標準偏差は2-2.5の区間で他のものとは違った数値を示している．これは編集量の標準偏差はほとんどの編集者が同じくらいであることを示している．

図 24　個人ページサイズと総編集回数の 両対数グラフ

この散布図では総編集回数が1000回くらいの人が多いことがわかり，その人たちの個人ページサイズは10000kb程であることを示している．

図 25　個人ページサイズと編集回数の両対数グラフ

この散布図では，1ヵ月間の編集回数が多い人も少ない人も個人ページサイズが100kb以上であることを示している.

## 5.3. 考察

Wikipediaでは利用者が編集した履歴を誰でも見る事ができ，自分のページを持っていれば，どんなことに興味があるのかなどを不特定多数の人に見て貰うことができる．多くWikipediaのページを編集してくれる人は，自己顕示欲の表れではないかと考えられる．

貢献度合いを可視化するランキングや，貢献者についての情報を記録する仕組みの存在が，Wikipediaのようなオープンなプロジェクトの成功に寄与していることが示唆される．

# 謝辞

本研究を進めるにあたり，ご指導を頂いた卒業論文指導教員の矢吹太朗准教授に感謝致します．また，日常の議論を通じて多くの知識や示唆を頂いた皆様に感謝します