

Zusammenhang zwischen sportlichem Erfolg und Darstellung bei Wikipedia im Vergleich zwischen Sportlern und Sportlerinnen

Lisa Hillebrand¹, Josefine Busch¹, Daniela Tumbrägel¹, and Flip Jansen¹

¹HTW Berlin

January 13, 2019

Abstract

tle—

1 Einleitung, Problemformulierung // Tumbrägel, Daniela (558529)

Die vorliegende Studie ist eine explanative Analyse des Zusammenhangs zwischen sportlichem Erfolg von Sportler*innen und dessen Repräsentation auf Wikipedia, gemessen an der Länge der Artikel und der Anzahl der Editierungen. Wie man am Beispiel von Caster Semenya feststellen muss, sind die Olympischen Spiele noch sehr von binären Geschlechterkonstrukten geprägt und unterscheiden die Teilnehmer*innen in ihren Wettbewerben dementsprechend (s. [IOC rules transgender athletes can take part in Olympics without surgery \(n.d.\)](#)). Daher ist die vorliegende Studie zwangsläufig dieser limitierten Sichtweise angepasst und unterscheidet nur zwischen den Kategorien Frauen und Männern.

Die Motivation für die Studie ergab sich aus dem Fall von Donna Strickland, die im Jahr 2018 - zusammen mit Gérard Mourou und Arthur Ashkin - einen Nobelpreis im Bereich der Laserphysik erhalten hat. Sie ist die dritte Frau, der je ein Physik-Nobelpreis verliehen wurde, und erst nach der Anerkennung wurde ein Wikipedia-Artikel über sie veröffentlicht (s. [Perspective — Why Nobel winner Donna Strickland didn't have a Wikipedia page \(n.d.\)](#)). Ihre männlichen Kollegen sind in der Online-Enzyklopädie schon seit mehreren Jahren vertreten. Einige Monate zuvor wurde ein Eintrag über Strickland abgelehnt, da sie von einem der Wikipedia-Editoren als nicht wichtig genug erachtet wurde (s. [Draft:Donna Strickland - Wikipedia \(n.d.\)](#)). Neben zahlreichen anderen Beispielen zeigt dies die geschlechtsspezifische Diskriminierung auf Wikipedia und insbesondere die Ausgrenzung von Frauen in der Wissenschaft auf (s. [Female Nobel prize winner deemed not important enough for Wikipedia entry \(n.d.\)](#)).

Geplant war, durch unsere Untersuchung zu überprüfen, ob Frauen in der Wissenschaft auf Wikipedia anders repräsentiert werden als Männer. Bei der Suche nach effizienten Datenquellen und Datensätzen ergab sich allerdings die Problematik, dass Frauen in der Wissenschaft insgesamt weniger vertreten sind und es somit sehr schwierig ist, einen sinnvollen Vergleich zwischen Frauen und Männern herzustellen. Deshalb ist die Entscheidung für einen anderen Themenbereich gefallen, in dem Frauen und Männer vergleichbaren Maßen

vertreten sind: Sport. Hier ist es möglich ein klares Ranking von individuellen Leistungen aufzustellen und Teilnehmer*innen auf dieser Grundlage zu vergleichen.

”Research shows just 16% of Wikipedia editors are female and only 17% of entries dedicated to notable people are for women”([Wikipedia biases](#) (n.d.))

Die Online-Enzyklopädie Wikipedia wird immer wieder für diverse Vorurteile und Tendenzen kritisiert. Darunter fallen Kategorien wie Gender, Kultur und Sprache sowie Geschichte und Politik (s. [Wikipedia biases](#) (n.d.)). Diese Studie konzentriert sich dabei auf die Kategorie Gender und untersucht diese im Fokus der Olympischen Sommerspiele 2016 in Rio de Janeiro.

Im Hinblick auf die größtenteils gleichmäßige Anzahl von Frauen und Männern wird das Verhältnis von sportlichem Erfolg und der entsprechenden Darstellung auf Wikipedia untersucht. Motiviert durch die oben genannten Ereignisse, bildet sich unsere These, dass Frauen bei gleichzusetzender Leistung im Vergleich zu Männern auf Wikipedia unzureichend repräsentiert werden. Im Detail wird dabei die mittlere Länge und die mittlere Anzahl der redaktionellen Änderungen der Wikipedia-Artikel der Sportler*innen differenziert.

Theorie

Wikipedia

Wikipedia wurde im Jahr 2001 gegründet ([Wikipedia Timeline](#), n.d.). Mittlerweile gibt es sprachenübergreifend über 46 Millionen Artikel [Wikipedia - Size of Wikipedia](#) (n.d.). Von rund 35 Millionen registrierten Nutzer:innen sind jedoch nur etwa 124.000 monatlich aktiv [Wikipedia - Wikipedians](#) (n.d.). Wikipedia ist zum Zeitpunkt des Verfassens dieses Texts auf dem fünften Platz der weltweit populärsten Websites [Alexa Top 500 Global Sites](#) (n.d.) und hat damit eine enorme Reichweite. Jede Person kann zu Wikipedia beitragen und auch ohne Registrierung Artikel bearbeiten [Wikipedia - Tutorial/Registration](#) (n.d.). Das Erstellen neuer Artikel ist nur nach vorheriger Registrierung möglich. Jeder Artikel verfügt über eine Versionsgeschichte, in der vorherige Artikelversionen sowie die Anmerkungen der Autor:innen und Administrator:innen gespeichert werden.

Gender Bias

Als Gender Bias bezeichnet man eine Verzerrung aufgrund der Geschlechtszugehörigkeit (QUOTE).

Gender Bias und Wikipedia

In mehreren Studien konnte gezeigt werden, dass ein Gender Bias sowohl hinsichtlich der Zusammensetzung der Autorenschaft bei Wikipedia als auch hinsichtlich der inhaltlichen Gestaltung der Artikel besteht. In der Autorenschaft sind bestimmte Personengruppen deutlich unterrepräsentiert sind. In Nutzerbefragungen wurde erfasst, dass zwischen 84% und 90% aller Wikipedia-Autor:innen männlich sind [Wikimedia Foundation, Inc. Community Engagement Insights/2018 Report](#) (n.d.) ([Graells-Garrido, Lalmas, & Menczer, 2015](#)). [Wagner \(2015\)](#)

Weiterhin zeigt die Studie von [Graells-Garrido et al. \(2015\)](#), dass der Gender Bias auf Wikipedia mit Struktur und Inhalt von Artikeln zusammenhängt. So zeigte sich ein signifikanter Unterschied in der Artikellänge, wengleich die Effektstärke nur gering ausfiel. Auf inhaltlicher Ebene fanden die Autor:innen, dass

Hypothesen

Es ergeben sich folgende Hypothesen:

Hypothese 1. Wikipedia-Artikel über Sportler sind signifikant länger als Wikipedia-Artikel über Sportlerinnen.

Hypothese 2. Wikipedia-Artikel über Sportler wurden signifikant häufiger editiert als Wikipedia-Artikel über Sportlerinnen.

Methodik // Busch, Josefine 560106

Erhebungsart

Zur Beschaffung der Daten stehen uns zwei Möglichkeiten zur Verfügung. Auf der offiziellen Webseite der Olympischen Kommission sind die Namen der Teilnehmenden der vergangenen Spiele veröffentlicht. [Olympic Games — Winter Summer Past and Future Olympics \(n.d.\)](#). Der Vorteil dieser Daten ist, dass sie bereits in weibliche und männliche Teilnehmer*innen unterteilt sind. Alternativ sind auch direkt auf Wikipedia Listen der Teilnehmenden zu finden. [Category:Competitors at the 2016 Summer Olympics - Wikipedia \(n.d.\)](#) Diese Listen beinhalten direkte Verlinkungen zu den jeweiligen Seiten der Sportler*innen. Allerdings ist hier nicht mehr ohne Weiteres auszulesen, welchem Gender die Person angehört. Um diese Information zu bekommen müsste der Fließtext auf die verwendeten Pronomen untersucht werden. Auch andere Informationen sind bei Wikipedia nicht mehr so leicht auslesbar wie aus der Liste von [olympic.org](#). Besonders sind die Sportler*innen nach dem Alphabet geordnet nicht nach dem erreichten Rang, dieser müsste also ebenfalls aus dem Fließtext ausgelesen werden.

Die Liste der Olympischen Spiele 2016 von [olympic.org](#) enthält strukturierte Daten von denen uns speziell nur die Spalten "Sport", "Discipline", "Event", "Names" und "Gender" interessieren. Die Länge der Artikel, gemessen an der Anzahl der Wörter, und die Anzahl der Editierungen der einzelnen Artikel müssen erst berechnet beziehungsweise von Wikipedia direkt abgefragt werden. Diese Daten liegen in unstrukturierter Form vor.

Stichproben

Da nicht in allen Sportarten Frauen und Männer in vergleichbaren Disziplinen antreten und in manchen Disziplinen Teams gegeneinander antreten, wählen wir 8 Disziplinen aus unterschiedlichen Sportarten aus, welche unsere Anforderungen erfüllen. Die Anforderungen sind, dass die Disziplinen für Frauen und Männer gleich sind bzw. gleich bezeichnet sind und dass sie keine Teamdisziplinen sind.

Die ausgewählten Disziplinen sind:

- Turmspringen 10m
- Bogenschießen Einzelwettkampf
- Fechten Épée Einzelwettkampf
- Moderner Fünfkampf Einzelwettkampf
- Leichtathletik Stabhochsprung

- Schwimmen 100m Freistil
- Radsport Straße Einzelwettkampf
- Leichtathletik Sprint 100m

Um die Artikel von Wikipedia zu bekommen und von diesen die Länge in Wörtern verwenden wir ein Modul von GitHub Benutzer Jonathan Goldsmith [goldsmith/Wikipedia](#) (n.d.). Mit diesem Modul und den Titeln der Wikipedia-Seiten können die Artikel abgerufen werden. Danach muss nur noch die Anzahl der Wörter ermittelt werden. Mit Stichproben soll geprüft werden, dass die korrekten Artikel zurückgegeben und gezählt werden.

Für die Zwischenspeicherung der Daten im Jupyter Notebook verwenden wir *pandas* Dataframes. Außerhalb des Programms sind die Daten in CSV-Dateien abgelegt. Für die grafische Darstellung im Notebook verwenden wir außerdem *matplotlib.pyplot* und *seaborn*.

Auswertungsmethoden

Für die Darstellung und Auswertung der Daten unterscheiden wir nach Gender der Sportler*innen und den Disziplinen.

Verteilung der Daten

Um ein besseres Verständnis für unsere Daten zu entwickeln wollen wir uns diese zunächst in einem Histogramm anschauen. Das Histogramm trägt auf der X-Achse die Anzahl der Wörter beziehungsweise Editierungen auf und auf der Y-Achse die Häufigkeit, mit der die jeweilige Anzahl vorkam. Dies soll uns zeigen wie die Daten verteilt sind und ob eine Normalverteilung vorliegt.

Mit weiteren Balkendiagrammen wollen wir die Extrem-, Mittel- und Medianwerte darstellen.

Ausreißer

Wir gehen davon aus, dass wir einige Ausreißer in unseren Daten haben werden. Sportler wie beispielsweise Usain Bolt werden erwartbar vergleichsweise lange Wikipedia-Artikel aufweisen. Um einen besseren Überblick über die Anzahl und Verteilung der Ausreißer zu bekommen, werden wir die Daten in einem Boxplot und einem Stripplot darstellen. Der Boxplot zeigt außerdem die Quantile. Als Ausreißer werden alle Datenpunkte im Boxplot dargestellt, die außerhalb des Bereiches liegen, in dem 90% der Daten zu finden sind.

Signifikanz

Sollten unsere Daten normalverteilt sein, werden wir die Signifikanz mit Hilfe des T-Tests untersuchen, andernfalls wollen wir den Mann-Whitney-U-Test verwenden.

”The Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed.”

2 Durchführung // Jansen, Flip 558059

2.1 Datenbeschaffung und Datenpräparation

Ausgehend von der englischsprachigen Wikipedia-Übersichtsseite der Teilnehmer*innen der Olympischen Sommerspiele 2016 *Wikipedia Category: Competitors at the 2016 Summer Olympics* (n.d.) haben wir rekursiv alle Sportarten gesammelt. Nachfolgend haben wir über die Sportarten die URLs und damit die Wikipedia-Titel der einzelnen Sportler*innen einer Disziplin gesammelt.

Für das Verarbeiten der HTML-Seiten haben wir die Python-Bibliothek *lxml lxml - Processing XML and HTML with Python* (n.d.) verwendet.

Um mögliche Fehler beim Sammeln der Daten von *olympic.org* zu vermeiden, haben wir das Olympic Studies Centre per Mail kontaktiert. Diese haben uns daraufhin eine Excel-Tabelle mit den ausführlichen Ergebnissen der Olympischen Sommerspiele 2016 zur Verfügung gestellt.

Die Daten haben wir anschließend nach den von uns gewählten Stichproben gefiltert, sodass daraus 8 Disziplinen mit den jeweils am besten platzierten 30 Frauen und 30 Männern resultierten.

Um die Wikipedia-Titel mit den von uns benötigten Daten der Ergebnistabelle zu verbinden, mussten wir nun die unterschiedlichen Namensformate anpassen und Sonderzeichen umwandeln. So setzt zum Beispiel Wikipedia bei Namensgleichheit von in Wikipedia erfassten Personen eine Kategoriebezeichnung ans Ende des Namens, z.B. "(diver)".

Trotzdem konnten wir einige URLs nicht automatisiert zuordnen, da es zum einen in der Ergebnistabelle Rechtschreibfehler in den Namen gab, zum anderen bei Namen aus Sprachgebieten mit nicht-lateinischer Schrift unterschiedliche Umschriften möglich sind. Auch wurde die Reihenfolge von Namensbestandteilen insbesondere aus dem asiatischen Raum nicht einheitlich gehandhabt. Teilweise ließ sich dieses Problem dadurch lösen, die Namensbestandteile permutierend zu vergleichen.

Ein Versuch, möglichen Rechtschreibfehlern bzw. Problemen mit Sonderzeichen zu begegnen, haben wir mithilfe des Levenshtein-Algorithmus *Levenshtein-Distanz – Wikipedia* (n.d.) unternommen. Dieser gibt die Anzahl von Einfüge-, Lösch- und Ersetz-Operationen zurück, die notwendig sind, um eine Zeichenkette in eine zweite umzuwandeln. Trotz niedrig angesetzten Schwellenwerten gab es aber zu viele falsche Zuordnungen, sodass wir darauf verzichtet haben. Für eine umfangreichere Datenauswertung lohnt es sich vermutlich, weiter in diese Richtung zu forschen.

Schlussendlich mussten wir einige Wikipedia-URLs der Sportler*innen manuell den entsprechenden Namen zuordnen.

Mit den bereinigten Daten konnten wir nun mit Hilfe der Wikipedia-Bibliothek für Python *goldsmith/Wikipedia* (n.d.) den zum Titel gehörenden Artikel auslesen. Nach Bereinigung des Artikels von Steuerungszeichen haben wir die Wörter gezählt und unserem *pandas* Dataframe hinzugefügt.

Die Anzahl der Editierungen konnten wir auf der "Page information Seite des jeweiligen Hauptartikels auslesen, die diesem Format folgt: <https://en.wikipedia.org/w/index.php?title=PYTHON&action=info>

Mit manuellen Stichproben haben wir die Richtigkeit unserer Daten überprüft.

3 Ergebnisse //Jansen, Flip 560106

Anzahl der Wörter

Verteilung

Bei der Betrachtung der Histogramme fällt sofort auf, dass es sich nicht um eine Normalverteilung handelt.

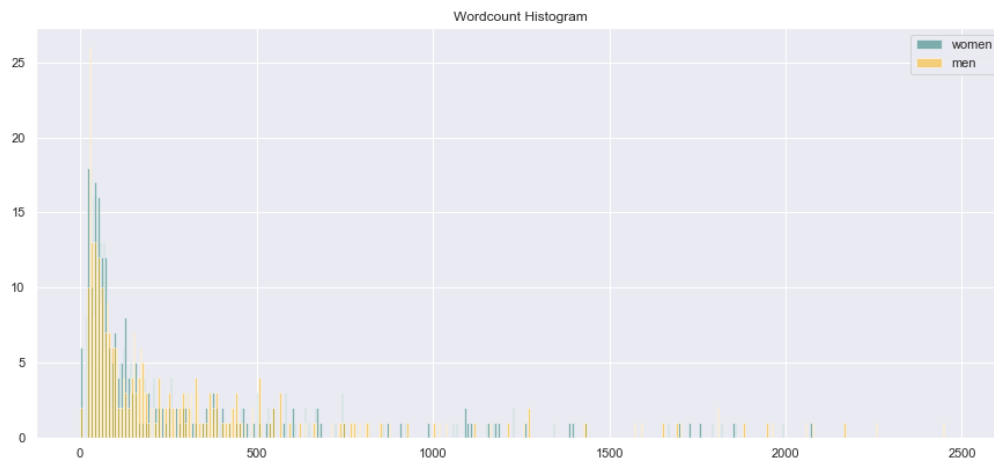


Figure 1: Histogramm der Wortanzahl über alle Disziplinen

Dieser Eindruck wird bekräftigt nach der Betrachtung der Histogramme für die unterschiedlichen Disziplinen.

Bei den Maxima fällt auf, dass es teilweise bedeutende Unterschiede gibt, sowohl zwischen den verschiedenen Sportarten als auch innerhalb der einzelnen Sportarten zwischen Männern und Frauen.

Die Artikel mit der höchsten Wortanzahl der drei Disziplinen Leichtathletik Sprint 100m, Radsport Straße Einzelwettkampf und Turmspringen 10m handeln von männlichen Sportlern. Es folgen mit geringerer Wortzahl Artikel über weibliche Sportlerinnen in den Disziplinen Radsport Straße Einzelwettkampf, Fechten Épée Einzelwettkampf und Leichtathletik Sprint 100m.

In 5 von 8 Disziplinen ist der Artikel mit der höchsten Wortanzahl über einen Mann, davon in 2 Disziplinen mit einem Vielfachen der Wortanzahl. In 3 Disziplinen ist der Artikel mit der höchsten Wortanzahl über eine Frau, davon ebenfalls in 2 Disziplinen mit einem Vielfachen der Wortanzahl.



Figure 2: Histogramme der Wortanzahl für alle Disziplinen

Disziplin	Maximum	Minimum	Mittelwert	Median	Standard Aweichung
Turmspringen 10m	1.690800e-01	nicht signifikant	1	1	1
Bogenschießen Einzelwettkampf	1.690800e-01	nicht signifikant	1	1	1
Fechten Épée Einzelwettkampf	1.690800e-01	nicht signifikant	1	1	1
Moderner Fünfkampf Einzelwettkampf	1.690800e-01	nicht signifikant	1	1	1
Leichtathletik Stabhochsprung	1.690800e-01	nicht signifikant	1	1	1
Schwimmen 100m Freistil	1.690800e-01	nicht signifikant	1	1	1
Radsport Straße Einzelwettkampf	1.690800e-01	nicht signifikant	1	1	1
Leichtathletik Sprint 100m	1.690800e-01	nicht signifikant	1	1	1

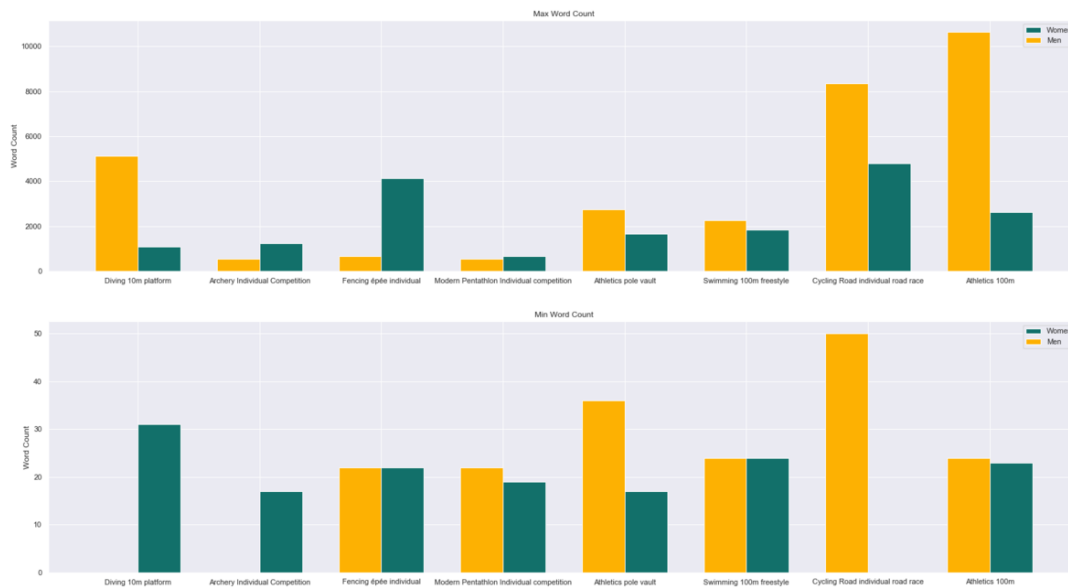


Figure 3: Maximum und Minimum der Disziplinen

Die Mittelwerte und Median-Werte der einzelnen Sportarten geben mit Ausnahme der Disziplin Radsport Straße Einzelwettkampf, bei der sowohl bei Mittelwert als auch Median die Artikellänge von männlichen die der weiblichen Sportler*innen um ein Vielfaches überschreitet, ein ausgeglicheneres Bild zwischen der Artikellänge von Männern und Frauen.

Auffällig ist, dass sich im Vergleich von Mittelwert und Median die Verhältnisse in 5 von 8 Fällen umkehren. Davon wird in 4 Fällen im Median die Artikellänge der Männer höher, wo sie im Mittelwert niedriger ist, und in einem Fall umgekehrt (Turmspringen 10m).

Das könnte auf Ausreißer verweisen, die im Median weniger Bedeutung haben.

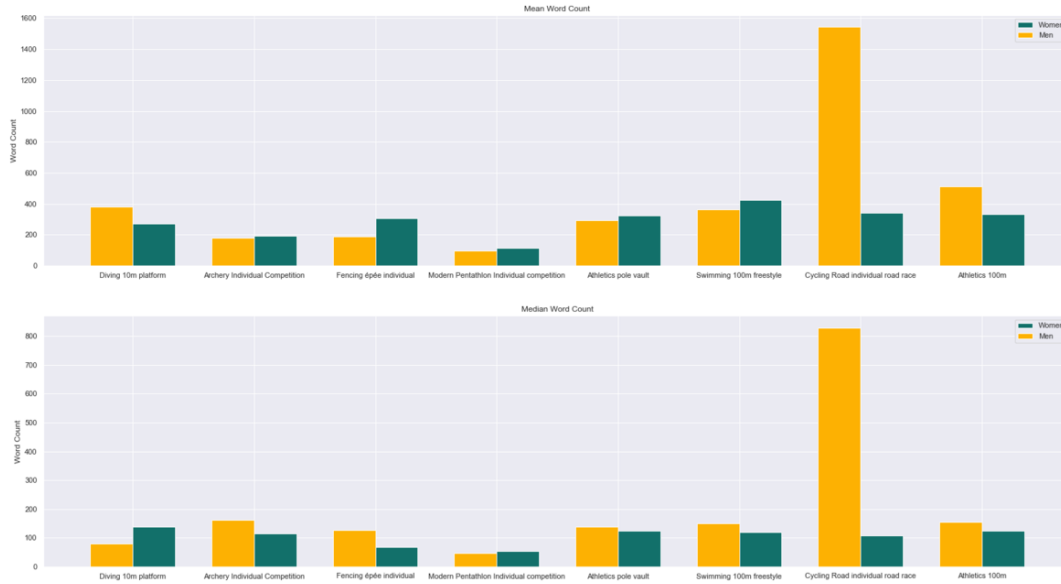


Figure 4: Mittelwert und Median der Disziplinen

Ausreißer

Anhand der Boxplots lässt sich zunächst feststellen, dass die Artikellänge der Ausreißer so markant ist, dass die Boxplots mit Ausnahme der Disziplin Radsport Straße Einzelwettkampf der Männer in der Darstellung fast verschwinden.

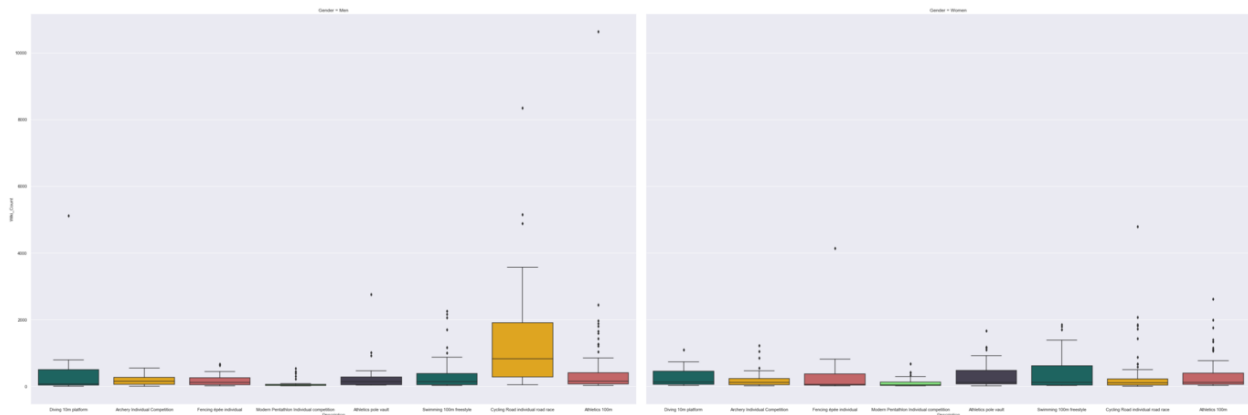


Figure 5: Boxplot der Disziplinen mit Ausreißern

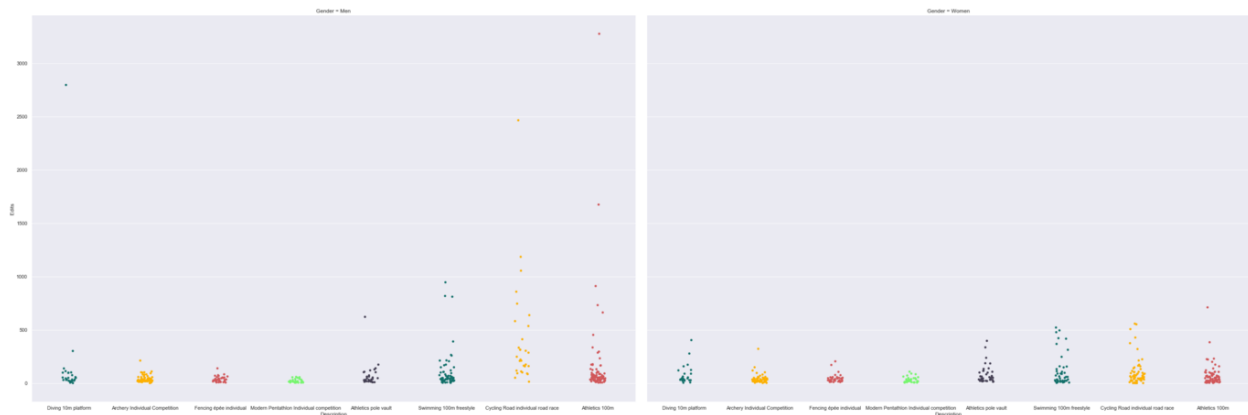


Figure 6: Stripplot der Disziplinen mit Ausreißern

Signifikanz

Disziplin	P-Value	Signifikanz
Turmspringen 10m	1.690800e-01	nicht signifikant
Bogenschießen Einzelwettkampf	2.440222e-01	nicht signifikant
Fechten Épée Einzelwettkampf	2.906772e-01	nicht signifikant
Moderner Fünfkampf Einzelwettkampf	3.782122e-01	nicht signifikant
Leichtathletik Stabhochsprung	2.451830e-01	nicht signifikant
Schwimmen 100m Freistil	4.563503e-01	nicht signifikant
Radsport Straße Einzelwettkampf	2.591859e-07	signifikant
Leichtathletik Sprint 100m	9.354463e-02	nicht signifikant

Wie aus den Median-Werten bereits erkennbar, sind nach Berechnung des P-Wertes die Unterschiede in der Artikellänge in 7 von 8 Disziplinen nicht als signifikant anzusehen. Lediglich in der Disziplin Radsport Straße Einzelwettkampf ist ein signifikanter Unterschied zugunsten der Artikellänge der männlichen Sportler*innen festzustellen.

Damit ist unsere Hypothese 1 widerlegt.

Anzahl der Editierungen // Daniela Tumbrägel (558529)

Verteilung

Wie bei der Anzahl der Wörter im zuvorkommenden Abschnitt fällt bei der Betrachtung des Histogramms auf, dass es sich nicht um eine Normalverteilung handelt.

Da die Histogramme der verschiedenen Disziplinen sehr linksbündig sind, wird diese Beobachtung dadurch zusätzlich bestätigt.

Die folgenden Grafiken zeigen jeweils die Maximal-, Mittel- und Medianwerte sowie die Standardabweichung der Editierungen pro Disziplin.

Mit Blick auf die Maxima der Editierungen per Disziplin ist erkennbar, dass es durchaus Unterschiede zwischen Frauen und Männern gibt, da insbesondere in den Disziplinen "Turmspringen 10m", "Radsport Straße

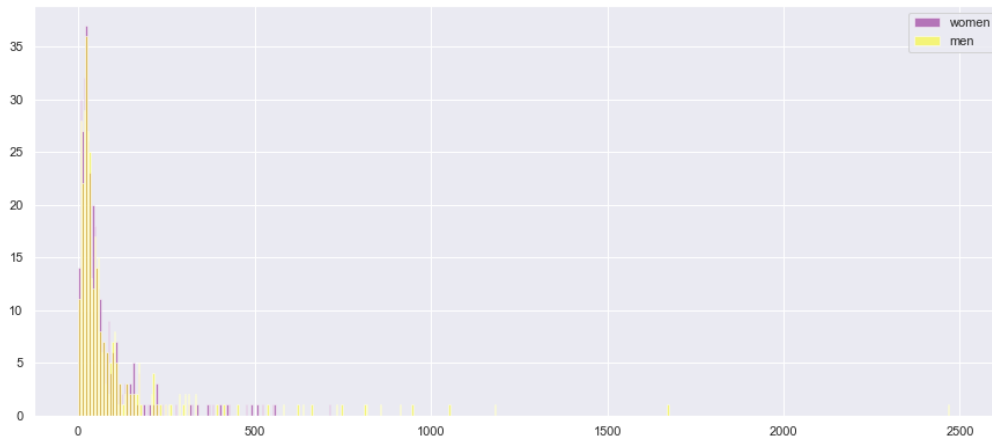


Figure 7: Histogramm der Editierungen über alle Disziplinen

Einzelwettkampfund "Leichtathletik Sprint 100m"bignifikante Unterschiede zugunsten der Männer sichtbar sind. Allerdings ist dies nur bedingt aussagekräftig, da es sich bei den Maxima um Ausreißer handeln kann, die nicht maßgeblich für die durchschnittliche Verteilung sind. Dies verifizieren wir durch nachfolgende Plots (s. Figure 10).

Das Histogramm, das den Mittelwert aufführt, zeigt ähnliche Beobachtungen, die jedoch ebenfalls aufgrund von Ausreißern schwer zu interpretieren sind.

Grundsätzlich kann aber festgehalten werden, dass sehr markante Extremwerte primär bei den Männern - mit einem Spitzenwert von 3276 Editierungen - zu finden sind.

Aufgrund der Verteilung der Daten ist der Median hingegen durch die Sortierung der Ergebnisse vergleichsweise aussagekräftig. Dieser visualisiert hier, dass es nur in der Disziplin "Radsport Straße Einzelwettkampf"bignifikante Unterschiede bei Frauen und Männern bzgl. der Editierungen der Wikipedia-Artikel gibt.

Disziplin	Maximum	Minimum	Mittelwert	Median	Standard Abweichung
Turmspringen 10m	1.690800e-01	nicht signifikant	1	1	1
Bogenschießen Einzelwettkampf	1.690800e-01	nicht signifikant	1	1	1
Fechten Épée Einzelwettkampf	1.690800e-01	nicht signifikant	1	1	1
Moderner Fünfkampf Einzelwettkampf	1.690800e-01	nicht signifikant	1	1	1
Leichtathletik Stabhochsprung	1.690800e-01	nicht signifikant	1	1	1
Schwimmen 100m Freistil	1.690800e-01	nicht signifikant	1	1	1
Radsport Straße Einzelwettkampf	1.690800e-01	nicht signifikant	1	1	1
Leichtathletik Sprint 100m	1.690800e-01	nicht signifikant	1	1	1

Ausreißer

Wie in der Methodik beschrieben, nutzen wir Box- und Stripplots insbesondere zur Visualisierung der Ausreißer in der vorhandenen Datenquelle. Die Abbildungen zeigen jeweils die Männer auf der linken und die

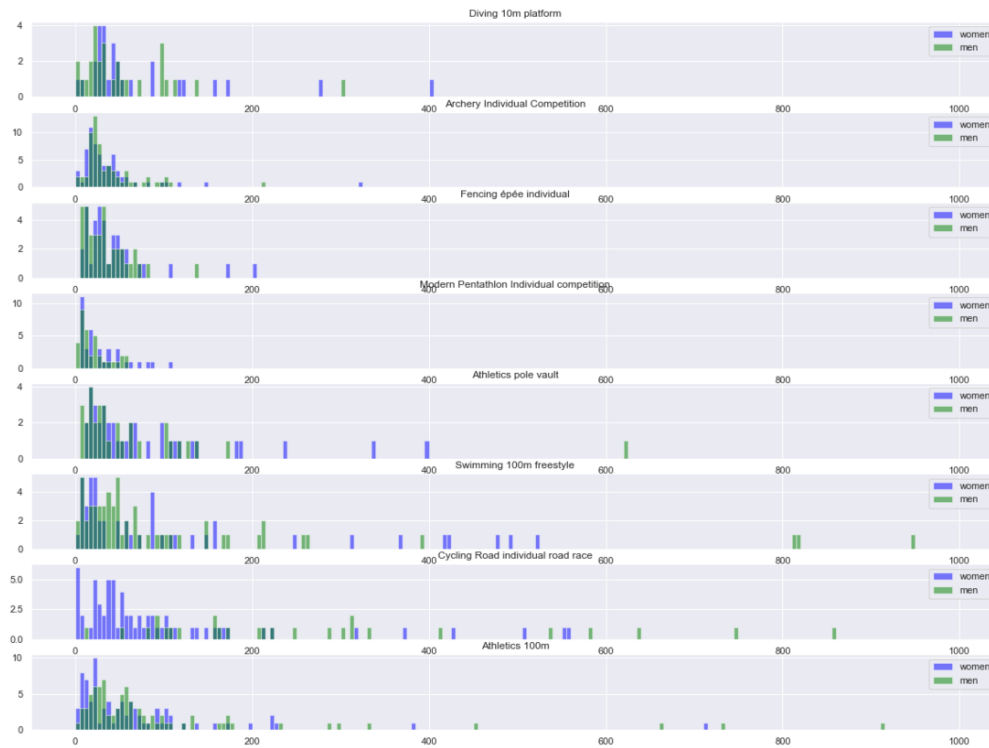


Figure 8: Histogramme der Editierungen für alle Disziplinen

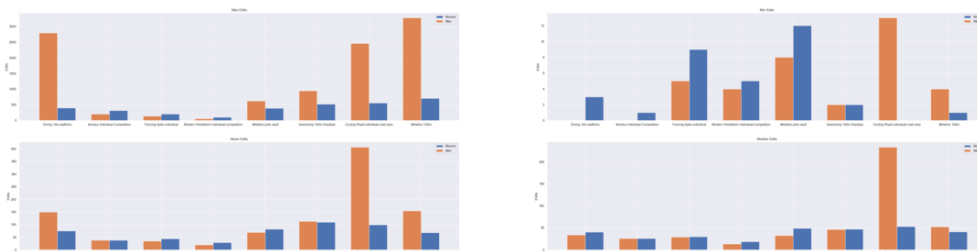


Figure 9: Maxima, Mean, Median und Standardabweichung der Editierungen

Frauen auf der rechten Seite.

Es ist deutlich erkennbar, dass bei den Männern mehrere sehr hohe Werte aufzufinden sind während es bei den Frauen zwar ebenfalls viele Ausreißer gibt, diese allerdings vergleichsweise niedrig angesiedelt sind. Die Disziplinen "Turmspringen 10m", "Radsport Straße Einzelwettkampfund "Leichtathletik Sprint

100m”fallen bei den Männern durch vereinzelte hochwertige Ergebnisse besonders auf. Bei den Frauen ist nur ein vergleichsweise hochwertiger Ausreißer in der Disziplin ”Leichtathletik Sprint 100m” zu erkennen, der sich allerdings nur marginal von den restlichen Ausreißern abgrenzt.

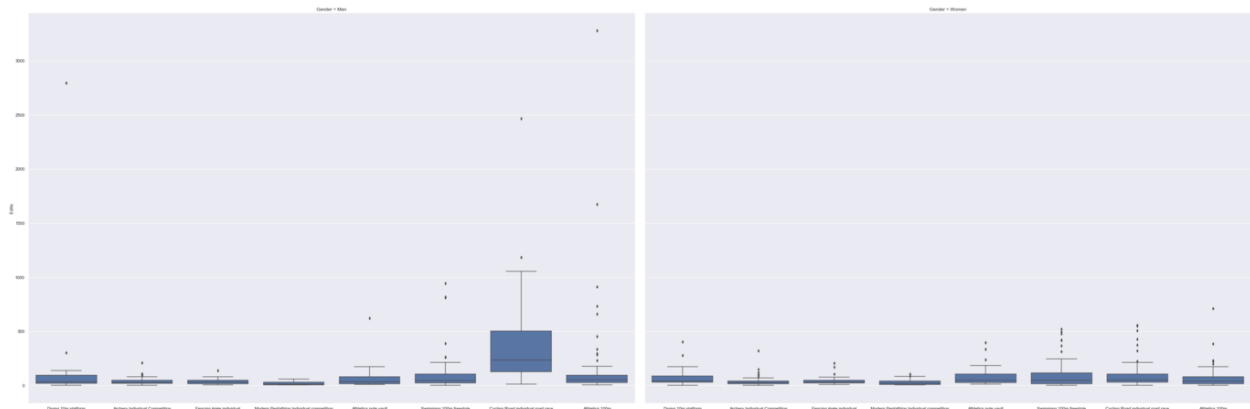


Figure 10: Boxplot der Disziplinen mit Ausreißern

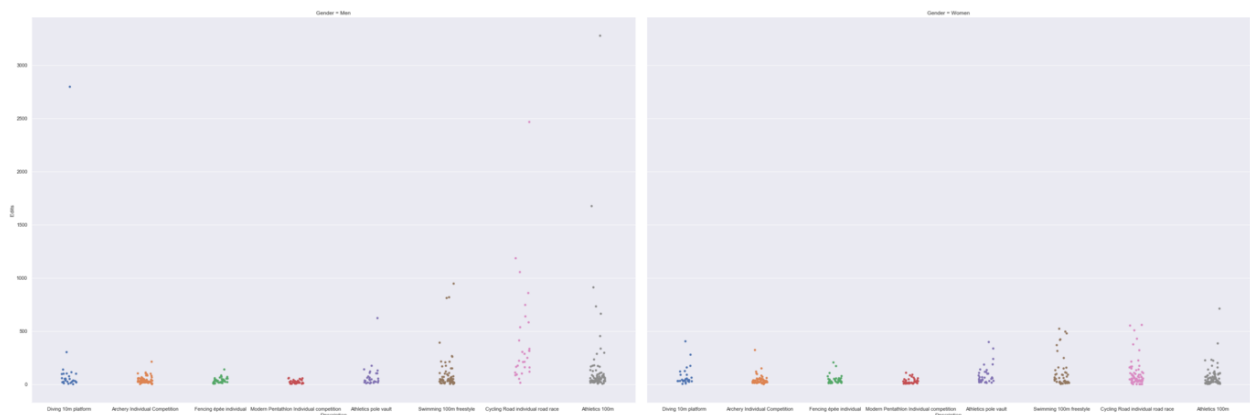


Figure 11: Stripplot der Disziplinen mit Ausreißern

Signifikanz

Der Mann-Whitney-U-Test zeigt mit Hilfe der Berechnung des P-Wertes, wie oben bei den Medianwerten bereits angedeutet, dass es größtenteils keine signifikanten Unterschiede zwischen Frauen und Männern gibt. Wie bei der Anzahl der Wörter ist auch bei der Anzahl der Editierungen in 7 von 8 Disziplinen keine Signifikanz zu erkennen.

Disziplin	P-Value	Signifikanz
Turnspringen 10m	1.690800e-01	nicht signifikant
Bogenschießen Einzelwettkampf	2.440222e-01	nicht signifikant
Fechten Épée Einzelwettkampf	2.906772e-01	nicht signifikant
Moderner Fünfkampf Einzelwettkampf	3.782122e-01	nicht signifikant
Leichtathletik Stabhochsprung	2.451830e-01	nicht signifikant
Schwimmen 100m Freistil	4.563503e-01	nicht signifikant
Radsport Straße Einzelwettkampf	2.591859e-07	signifikant
Leichtathletik Sprint 100m	9.354463e-02	nicht signifikant

Die zweite Hypothese ist somit ebenfalls widerlegt.

4 Fazit / Diskussion

References

- Alexa Top 500 Global Sites*. (n.d.). <https://www.alexa.com/topsites>. Retrieved from <https://www.alexa.com/topsites> (Accessed on Tue, January 08, 2019)
- Category:Competitors at the 2016 Summer Olympics - Wikipedia*. (n.d.). https://en.wikipedia.org/wiki/Category:Competitors_at_the_2016_Summer_Olympics (Accessed on Mon, December 31, 2018)
- Draft:Donna Strickland - Wikipedia*. (n.d.). https://en.wikipedia.org/w/index.php?title=Draft:Donna_strickland&oldid=842614385. Retrieved from (Accessed on Fri, January 11, 2019)
- Female Nobel prize winner deemed not important enough for Wikipedia entry*. (n.d.). <https://www.theguardian.com/science/2018/oct/03/donna-strickland-nobel-physics-prize-wikipedia-denied>. Retrieved from <http://www.theguardian.com/science/2018/oct/03/donna-strickland-nobel-physics-prize-wikipedia-denied> (Accessed on Fri, January 11, 2019)
- goldsmith/Wikipedia*. (n.d.). <https://github.com/goldsmith/Wikipedia>. Retrieved from <https://github.com/goldsmith/Wikipedia> (Accessed on Mon, December 31, 2018)
- Graells-Garrido, E., Lalmas, M., & Menczer, F. (2015). First Women Second Sex. In *Proceedings of the 26th ACM conference on hypertext & social media - HT '15*. ACM Press. Retrieved from <https://doi.org/10.1145/2700171.2791036> doi: 10.1145/2700171.2791036
- IOC rules transgender athletes can take part in Olympics without surgery*. (n.d.). <https://www.theguardian.com/sport/2016/jan/25/ioc-rules-transgender-athletes-can-take-part-in-olympics-without-surgery>. Retrieved from <http://www.theguardian.com/sport/2016/jan/25/ioc-rules-transgender-athletes-can-take-part-in-olympics-without-surgery> (Accessed on Fri, January 11, 2019)
- Levenshtein-Distanz – Wikipedia*. (n.d.). <https://de.wikipedia.org/wiki/Levenshtein-Distanz>. Retrieved from <https://de.wikipedia.org/wiki/Levenshtein-Distanz> (Accessed on Wed, January 09, 2019)
- lxml - Processing XML and HTML with Python*. (n.d.). <https://lxml.de>. Retrieved from <https://lxml.de> (Accessed on Wed, January 09, 2019)
- Olympic Games — Winter Summer Past and Future Olympics*. (n.d.). <https://www.olympic.org/olympic-games>. Retrieved from <https://www.olympic.org/olympic-games> (Accessed on Mon, December 31, 2018)
- Perspective — Why Nobel winner Donna Strickland didn't have a Wikipedia page*. (n.d.). https://www.washingtonpost.com/nobel-winner-donna-strickland-didnt-have-wikipedia-page/?noredirect=on&utm_term=.7b40ea4f5a11. Retrieved from (Accessed on Fri, January 11, 2019)
- Wagner, e. a., Claudia. (2015). It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia. *ICWSM*.
- Wikimedia Foundation, Inc. Community Engagement Insights/2018 Report*. (n.d.). https://meta.wikimedia.org/wiki/Community_Engagement_Insights/2018_Report (Accessed on Sun, December 09, 2018)
- Wikipedia biases*. (n.d.). <https://www.theguardian.com/technology/2018/jul/29/the-five-wikipedia-biases-pro-western-male-dominated>. Retrieved from <http://www.theguardian.com/technology/2018/>

[jul/29/the-five-wikipedia-biases-pro-western-male-dominated](#) (Accessed on Fri, January 11, 2019)

Wikipedia Category: Competitors at the 2016 Summer Olympics. (n.d.). <https://en.wikipedia.org/wiki/Category:Competitors> (Accessed on Sat, January 05, 2019)

Wikipedia - Size of Wikipedia. (n.d.). https://en.wikipedia.org/wiki/Wikipedia:Size_of_wikipedia. Retrieved from (Accessed on Sun, January 06, 2019)

Wikipedia Timeline. (n.d.). https://meta.wikimedia.org/wiki/Wikipedia_timeline. Retrieved from (Accessed on Sat, January 05, 2019)

Wikipedia - Tutorial/Registration. (n.d.). <https://en.wikipedia.org/wiki/Wikipedia:Tutorial/Registration>. Retrieved from <https://en.wikipedia.org/wiki/Wikipedia:Tutorial/Registration> (Accessed on Sun, January 06, 2019)

Wikipedia - Wikipedians. (n.d.). <https://en.wikipedia.org/wiki/Wikipedia:Wikipedians>. Retrieved from <https://en.wikipedia.org/wiki/Wikipedia:Wikipedians> (Accessed on Sun, January 06, 2019)