Age, Gender, Nationality: How do they affect Twitter Circadian Rhythms?

Abstract

Different systems of the body follow circadian rhythms that are synchronized with a master clock in the brain which resets every 24hrs. Our study aims to address how socio-demographic elements like age, sex and nationality alter this master clock. We analyze a range of exogenous factors like reduced social interaction with age, work-life balance patterns across different countries and gender roles impacting sleep schedules. To render this idea statistically, we observe the amount of sleep that twitter users receive, on the basis of their age, gender and nationality. We are operating on around 1050000 tweets. We infer age and gender from profile pictures of Twitter users, and manually annotate their locations to nationality. Additionally, we use the DeepExtraction Model for inference and only pick rows having a high confidence. We analyze how each feature and their combinations affect the amount of sleep that these demographic groups get, paying special attention to mitigating any inherent bias that might be present. Further, we study why each of these factors affect sleep in the way they do.

1 Related Work

According to Wikipedia, a circadian rhythm, or circadian cycle, is a natural, internal process that regulates the sleep–wake cycle and repeats roughly every 24 hours. This natural process is sometimes referred to as the master clock, which is different from the biological clock. This master clock is a function of environmental factors, especially light, which is why circadian rhythms are tightly tied to the cycle of day and night. Other factors like jet lag, sleep disorder, irregular work shifts across different professions have been attributed as some of the causes of influencing sleep wake cycles. With the rising ease of accessibility to geolocated data, studies associating human digital foot-print with their daytime activities has received a volume of attention. Such research usually lever-age the hyper-connected social systems and how people mirror their offline self, online [7]. H.Schwartz et. al proposed how twitter vocab maybe used to predict a dynamic range of health, well-being, demographic and socioeconomic conditions. For instance, spotting excessive alcohol consumption [4], analysing the happiness-levels of people who have richer neighbours [6], predicting heart disease mortality as a function of negative emotions (especially, anger) [5]. This motivates us to leverage Twitter data to try and relate circadian rhythms of various demographic clusters to Twitter user activity.Intuitively, tweets do not seem to be the perfect method of estimating sleep wake cycle, however, even though it takes a brief moment to type 280 characters, tweeting is a definite record of wakefulness. Sustained periods of low Twitter activity have proven to be correlated with sufficient sleep[13], as found out through surveys. Inactivity duration on Twitter has been studied to relate late night tweeting and its effect on the performance of basketball players [9]. In essence, we see a bulk of studies coupling circadian rhythms with professions, daylight savings, weekend or weekday activities, however, our approach of incorporating socio-demographic factors is relatively novel. We base our approach on three major features, age, sex and nationality. A number of factors like social interactions, screen time and work-related physical activity reduce with the increase in age. Further, there seems to be a difference in the circadian rhythms of men and women because of hormonal differences, gender roles and how they report[14]. In [11] societies with access to limited electricity, activity is timed according to availability of daylight. In another wise modern society where electricity is avail-able in plenty, people live out of synchrony with the day-night cycle. Since occupation and economy comprise the major chunks of nationality, it makes sense to use this as one of our features.

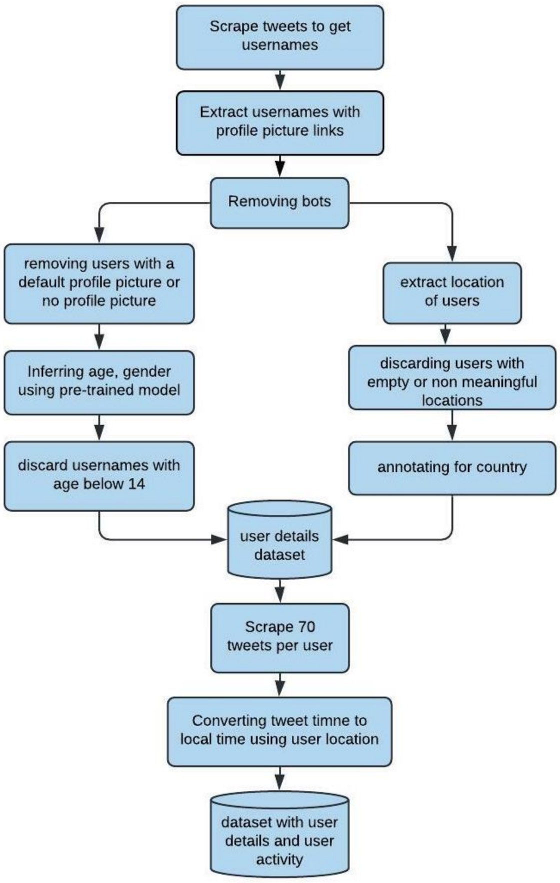

Figure 1: Flow chart for dataset generation

2 Data

We are particularly interested in assessing correlations between actual circadian rhythms with Twitter bedtime patterns. To achieve this, we assess data across age, gender and nationality. The dataset provided to us for the course project has user-names hashed because of which we are unable to infer age gender. Hence, for these two fields, we predict using a pre-trained CNN model that takes profile image as input and predicts the age and gender of the individual as output. The first step was to scrape the data using Twitter API. We scraped 125000 tweets with user name, location, profile image url, creation time for each tweet. We also scraped a column that tells us whether the image is a default image or a profile image has been set. Out of these 125000 tweets, we found 122185 unique users. Out of these 122185 users we filtered for bots and organization pages. We use Twitter Bot or Not [1] to filter such users out. Finally, after this we remain with 75625 users.

[1]Github: https://github.com/scrapfishies/twitter-bot-detection

	Username	COUNTRY	age	pred_gender	date	time	hour
71	1Chelin	SOUTH AFRICA	30.0	female	2021-04-15	17:21:11	17
72	1Chelin	SOUTH AFRICA	30.0	female	2021-02-25	15:56:21	15
73	1Chelin	SOUTH AFRICA	30.0	female	2021-01-08	12:03:49	12
74	1Chelin	SOUTH AFRICA	30.0	female	2021-01-08	11:51:57	11
75	1Chelin	SOUTH AFRICA	30.0	female	2020-12-25	13:12:56	13
76	1Chelin	SOUTH AFRICA	30.0	female	2020-11-08	07:15:10	7
77	1Chelin	SOUTH AFRICA	30.0	female	2020-11-03	12:07:11	12
78	1Chelin	SOUTH AFRICA	30.0	female	2020-09-30	21:51:00	21
79	1Chelin	SOUTH AFRICA	30.0	female	2020-09-24	11:14:32	11
80	1Chelin	SOUTH AFRICA	30.0	female	2020-09-22	12:14:35	12

Figure 2: Final dataset with columns

Next from the 75625 users we filtered out users who do not have a profile picture or use a default profile picture. After this step, we remain with only 5749 unique users. Next we annotated nationality manually for 5749 users. We observed that a total of 1315 rows were empty and location and some of the users did not have a meaningful location. Finally, after annotating all meaningful locations we have a total of 2910 users. We observed that the majority of the valid user locations were from the USA, UK, Canada, India, Japan, South Africa, Nigeria and Kenya.Initially we were interested in analysing the twitter usage in the most populous countries. However, Twitter is banned in China, hence we had to skip that.

Further, Wikipedia states that Nigeria has an ongoing indefinite ban on Twitter starting June 2, 2021. On the contrary, we found 148 out of 2910 users from Nigeria. Next we scraped 140 tweets for each of these users from the period of November 2021. To build the final dataset we joined the user details with the tweet activity per user. For preprocessing, we converted the time zones from UTC to local time depending on the user location. Finally we have a dataset with the columns as seen in 2.

3 Approach

3.1 Inferring Age

The first step in our approach was to enrich the dataset with age data. We have used profile pictures of users to predict the age group they belong to. The state-of-the-art model for this is DEX which has been pretrained on IMDB-WIKI dataset on 500K images using an improvised VGG-16. This model has an accuracy of 64% for predicting the exact age group and 90% for predicting age group by 1 off. The model outputs one out of the 8 age

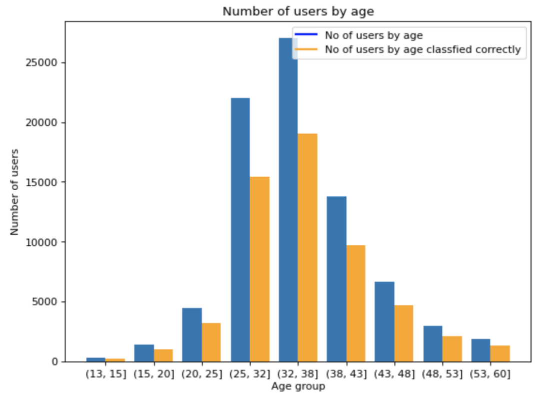

Figure 3: Bar chart for showing age distribution and prediction using a pre-trained model

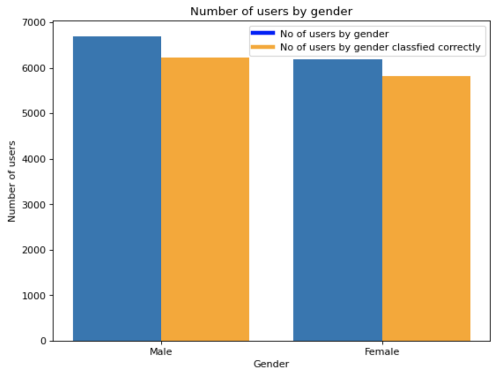

Figure 4: Bar graph for gender classification and prediction

groups (0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60- years). It makes more sense for us to use a categorical output as the sleep patterns are expected to be similar for an age group.

For preliminary observations, we tested this model on our dataset of 81220 twitter users. To generate this dataset we used profile image url https attribute of Twitter API user object. On passing the data through the DEX model, we achieved an average accuracy of 70% on the age categories mentioned above.

3.2 Inferring Gender

Gender is not an available field in Twitter API and hence it has to be inferred. For our approach we use the same model that we used for age inference, i.e., DEX. This model shows an accuracy of 90% for binary classification between male and female. We tested this model's gender prediction task on CrowdFlower's Twitter User Gender Classification dataset to acheive an accuracy of 93%. Finally, we used this model to predict age of users that we scarped. Since our feature is already a prediction, we would want to have a good accuracy on this field. To achieve this, we use the gender confidence field to filter rows that have a confidence of at least 99%.

3.3 Annotating Nationality

Only about 30% of twitter users put locations in their profiles. However, this amount is enough to determine the nationality of any Twitter user correctly. We study the tweet activities of the top 15 nations with the maximum number of twitter users. Due to unavailability of a dataset with tagged nationality, we decided to annotate the locations to their corresponding countries. Three annotators tagged a sample of 5000 random items (1000 samples 5 times) from the dataset. We removed rows without a valid location manually and annotated locations of the rest. Further we estimated unobserved (zero) counts. We preprocessed these taggings to prevent inconsistencies (for instance, annotator 1 wrote "USA" and annotator 2 wrote "United states of america").

4 Definitions

We observe that each daily activity has two peaks, one during the start of the day and the other is usually observed after the work day is over.

WakeTime: First hour of average tweet activity during the day (activity is 75% of the peak morning activity)

Sleeptime: Onset of 8 hour block of time with least activity (activity reduces to 25% of the peak activity in the evening)

5 Analysis

In this section, we try to understand the scale of our dataset in terms of age, gender, nationality and also analyse circadian rhythms for various groups. Every circadian rhythm plot has 24 data points, indicating the 24 hours of a day. The 24 hour period is an average of the Twitter activity of each user. More specifically, the 5th hour of the day in the plot is essentially a mean of the 5th hour of everyday. Note that we will use the terms described above in section 4.

5.1 Circadian Rhythm by Nationality

To study in depth, we analyse only 5 countries namely USA, UK, India, Japan and South Africa. Figure 6 lists the top 15 countries with highest

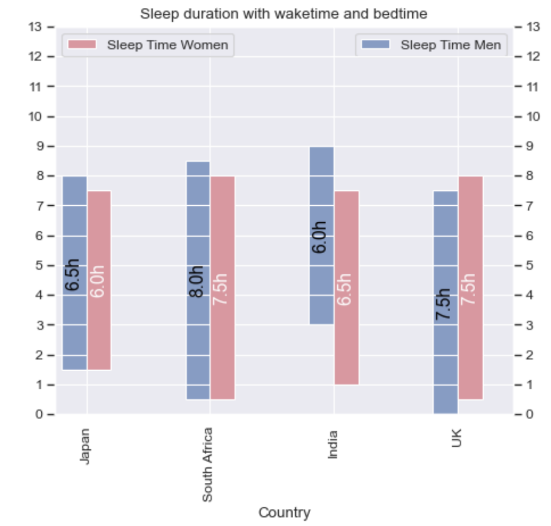

Figure 5: Gender wise sleep duration for 4 countries

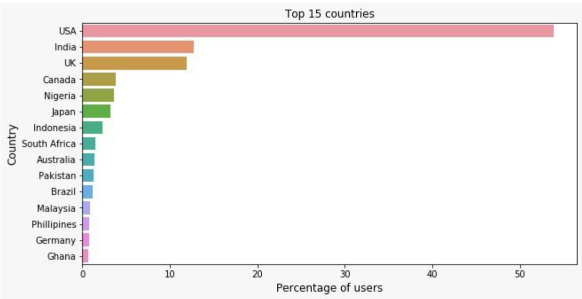

Figure 6: Percentage of users by country

tweet counts. To study the diversity in terms of timezones, economic levels and male-to-female ratio, we chose countries from major continents like America, Africa, Europe and Asia. This diversity will enable us to understand the effect of intertwining of age, gender and nationality in determining sleep cycles.

In figure 5, we visualise the waketime, bedtime and bedtime duration for four countries. The lower bound of each bar represents the bedtime whereas the upper bound represents the waketime.The difference between the two is interpreted as the sleep duration.We can observe that Japanese people go to sleep late and rise early making Japan a sleep deprived country. The next highly sleep deprived country in India.

5.1.1 Circadian Rhythm of Men and Women in South Africa

Before we analyze the circadian rhythm of South Africa it is important to note that it is the most industrialized, technologically advanced, and diversified economy in Africa [2]. We chose to omit Nigeria from our studies because of the indefinite conditional ban on Twitter, which could skew our results.

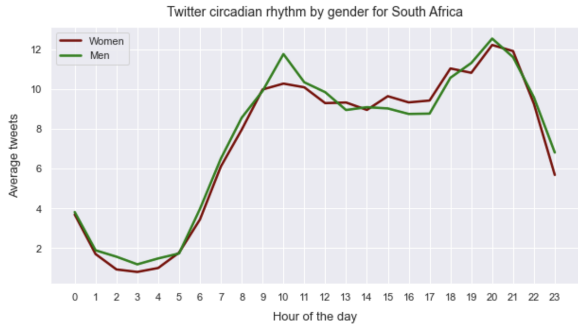

Figure 7: Twitter activity of men and women in South Africa

Twitter Waketime (where 75% of morning peak tweet activity happens) for women in South Africa is around 8AM whereas for men it is around 8:45 AM. Twitter Sleeptime for South African women is 12:30AM and for men is 12AM. A study conducted [2] to assess the sleep quality among diverse cultures claims that 31.3% of South African women and 27.2% of men reported difficulty with sleep. This is further validated in a study by NCBI [15] where they figured that sleep duration is 7.88 h among men and 7.46 h among women.

5.1.2 Circadian Rhythm of various age groups across UK

Sleep patterns were investigated [11] in a British population cohort to realise the effects of sleep in various health outcomes where sleep difficulties, specifically early awakening, were reported more often by women (14.7%) than men (11.7%). On average, ages between 18-29 in the UK get the most sleep with 51% of the cohort enjoying more than eight hours of sleep on average [3].

Figure 8 almost corroborates this study as it is evident that the age group of 14-24 have very low tweet activity all throughout the day, hence we cannot define waketime and sleep time for this group. The working class age group of 24-34 has the highest tweet activity, with 7% of the group having less than 6h of sleep. The twitter sleep duration for this group is 7h which can be explained by an extra hour of Time In Bed (TIB) described in [12].

[2] https://en.wikipedia.org/wiki/Economy_of_South_Africa

[3] https://www.firstbeat.com/en/blog/world-sleep-284 day-2020-how-much-sleep-do-different-age-285 groups-get-in-the-uk/

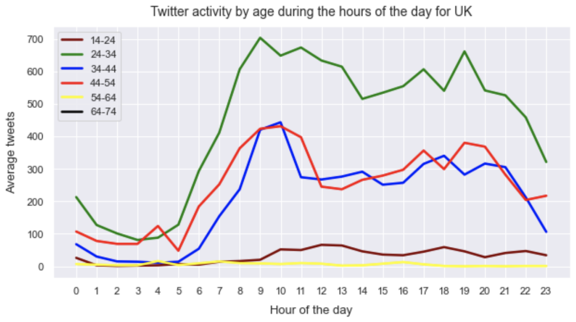

Figure 8: Twitter activity by average tweets for various age groups in UK

54+ is the age group, where Twitter activity fails to mimic the original circadian rhythm. This is because, although this age group gets the least sleep, they are also not quite acquainted with technology and hence Twitter usage is not a reliable marker of their wakefulness.

5.1.3 Circadian Rhythm of Japan

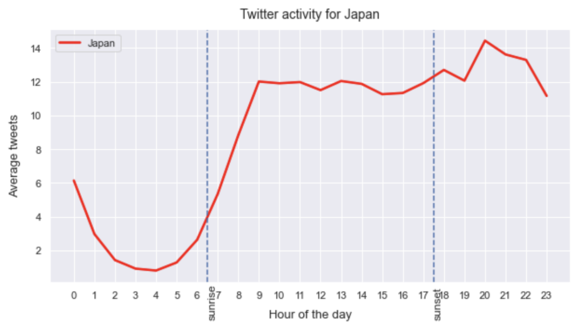

Figure 9: Circadian rhythm for Japan

Japanese people are the most sleep deprived with an average sleep duration of 6hr 15mins, and never goes up more than 7hr [4]. Their Twitter usage shows that their average sleep duration is around 6.5h. Their constant sleep deprivation might help explain why many Japanese people are found napping in public. [5]. Japan's work culture has been associated with stress [10] which is again a good indicator of how sleep deprivation is a lasting concern in Japan. Their sleep deprivation is evident from figure 9. In the month of November, i.e., when tweets were scraped, the sunrise time in Japan was around 6:50AM, 40mins after which Japanese women are active on Twitter. Japan has only half the population of the USA, and is the second largest market of Twitter with a 169 million dollars revenue. This validates the extensive amount of tweets and duration of Twitter activity in Japan.

Gender	WakeTime	BedTime	Sleep duration
Men	8 AM	1:30 AM	6 h 30 m
Women	7:30 AM	1:30 AM	6 h

Table 1: WakeTime and BedTimes for Japan by gender

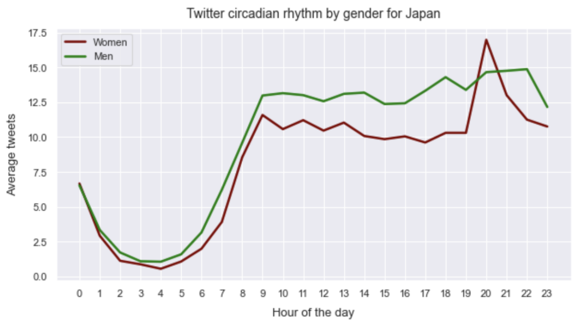

Figure 10: Daily Twitter usage based on gender for Japan

Figure 10 shows that there is a significant difference in the Twitter Sleep Cycle of men and women in Japan. The Gender Inequality Index (GII) Japan ranked as 19th out of 188 countries in 2019 which is also the worst among G7 countries. In 2019, 57.4 percent of all women participated in the labor force, however, traditional gender roles are still common due to the prevalence of patriarchal systems in Japan, which explains the lower sleep duration among women.

5.1.4 Circadian Rhythm of English and Non-English Countries

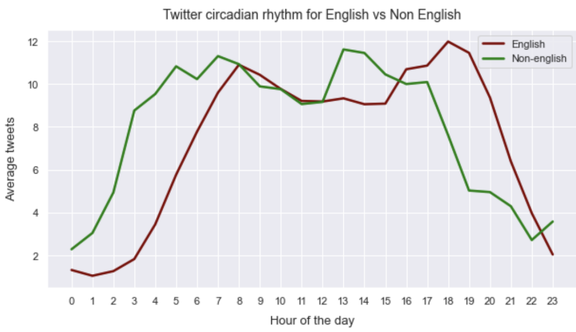

Figure 11: Circadian rhythm of English vs Non-English Speaking countries

USA, UK are grouped as English Speaking countries and South Africa, India and Japan are grouped as Non-English speaking countries. In figure 11,

[4]https://blog.withings.com/2014/11/05/cultural-differences-impact-on-sleep-patterns/

[5]https://www.nytimes.com/2016/12/16/world/what-in-the-world/japan-inemuri-public-sleeping.html

we observe that the Twitter activity of English speaking nations is shifted to the right as compared to Non-English speaking nations. South Africa wakes up the earliest, which could result in the shift of the plots. [6] The duration of Twitter activity in Non-English speaking countries is more than the English speaking ones. This is expected since India and Japan are the most sleep deprived countries in the world.

5.2 Circadian Rhythm by Age

Currently, we have divided the age of users into bins of size 10 (chosen arbitrarily) starting from 14 to 74.

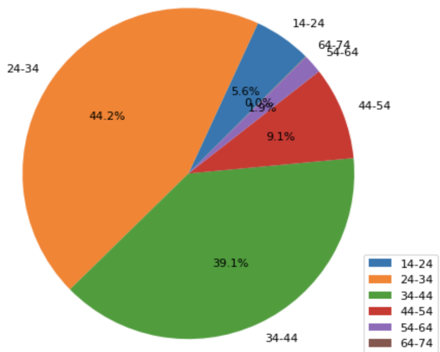

Figure 12: Distribution of Twitter users by age in the USA

Figure 12 shows close to 45% of the users in the USA lie in the age group of 24-34. This is followed by the age group 34-44 which has about 40% of users. These two age groups constitute 85% of the Twitter users. The chart has no sector for age group 64-74 because our dataset does not contain any such user.

Fig 13 demonstrates the Twitter activity of different age groups which, for an hour, is simply the number of tweets from users in that age group during that hour. At any given time the tweets from the age groups 34-44 and 24-34 is significantly larger than [7]other cohorts. Twitter activity for this age group reaches peak around 8AM and then plateaus till 6PM, after which there is a second peak. This can be attributed to the office hours of this age, since they represent the working population.

[8] Almost half of all Americans say they feel sleepy during the day between three and seven days per week [9] . The 24-44 cohort, with highest twitter activity, represents around 85% of the US population, which rightly explains the sleep deprivation that most Americans feel. Older adults are largely obscured in social media representations, hence, their twitter activity is not true representation of their circadian rhythm. Although this age cohort signed up for social media to reduce isolation [3], their activity remains low due to reduced social interaction with age.

Around 34% of Americans take a nap on a typical day [10]. 38% men and 31% women report that they caught a snooze in the past 24 hours. Figure 13 shows the dip in activity between 1PM and 4PM which may be realised as the naptime. However, this should be taken with a pinch of salt since the dip in Twitter usage of age cohort 24-44 could also indicate their work hours.

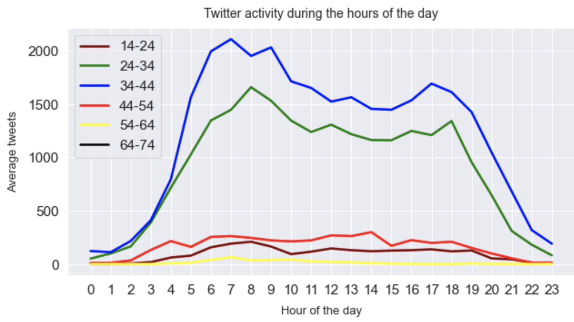

Figure 13: Twitter activity by age in USA

Age/Times	**24-34**	**34-44**
BedTime	11:30PM	11PM
WakeTime	6AM	6AM
Bedtime duration	10 hours	11 hours
Naptime	3 hours	2 hours

Table 2: WakeTimes and BedTimes of Twitter users by age

5.3 Circadian Rhythm by Gender

Figure 14 shows the distribution of males and females in five countries as mentioned in 5.1. USA, South Africa and UK have a high percentage of female users as opposed to other countries. Per 100 female, a study found that the male-to-female ratio of USA is 97.95, South Africa is 97.09 and UK is

[6]https://businesstech.co.za/news/lifestyle/86268/south-africans-are-the-earliest-risers-in-the-world/

[7]https://sproutsocial.com/insights/new-social-media-demographics/

[8]https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/

[9]https://www.sleepfoundation.org/wp-content/uploads/2020/03/SIA-2020-Q1-Report.pdf

[10]https://www.pewresearch.org/social-trends/2009/07/29/nap-time/

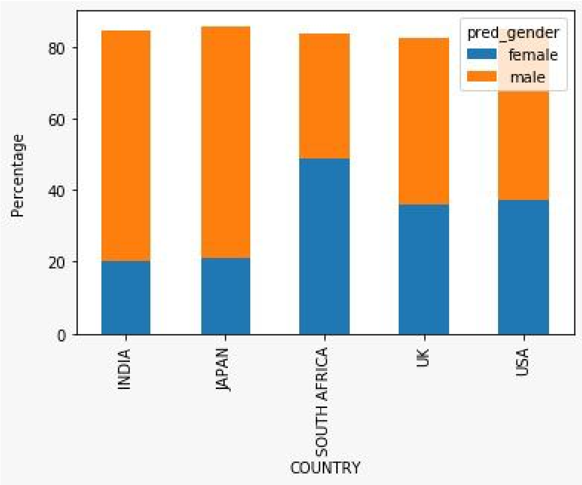

Figure 14: Twitter users per country by gender

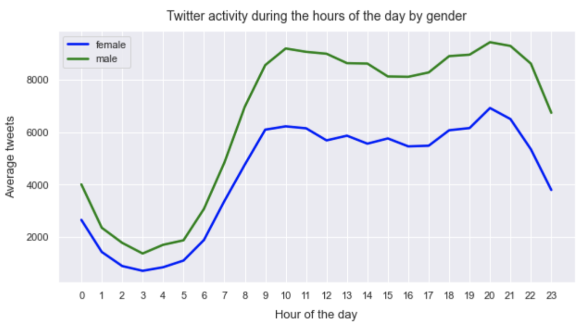

Figure 15: Twitter activity by gender

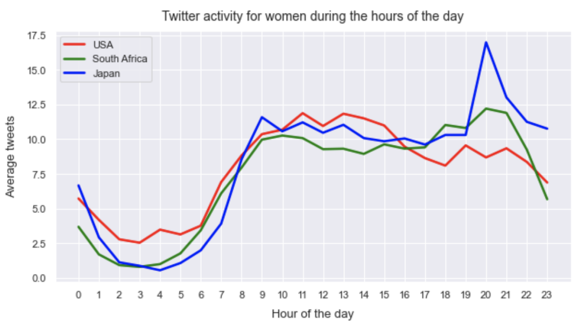

Figure 16: Twitter activity of women in different countries

97.67. We further observe that around 38% of the twitter users are female and the rest 62% are males.

In figure 15 we analyze the Twitter activity based on gender.

Using the previous mentioned 75%-25 heuristic, we have summarized Bedtimes and Waketimes in the Table 3. The table shows that women on an average sleep 30 minutes more than men as also shown by [11]this study.

[11]https://blog.withings.com/2014/11/04/study-of-the-sleep-patterns/

Gender/Times	**Men**	**Women**
BedTime	10PM	9PM
WakeTime	5:30AM	5AM
Bedtime duration	7:30 hours	8hours

Table 3: WakeTimes and BedTimes for men and women

Figure 16 represents the circadian rhythm of women is USA, South Africa and Japan. We can observe that the peak activity for USA is realised later than Japan and South Africa. This implies that women in Japan and South Africa start their day earlier than women in USA. The figure also shows that the Japanese women sleep later than the other women. This could be attributed to the fact that Japan has the most active twitter users, and also, its the most sleep deprived country in the world.

6 Evaluation

To evaluate the strength of our analysis, we have used Kolmogorov–Smirnov test [1]. The D statistic is the absolute maximum distance between the CDFs of two samples. The closer this number is to 0 the more likely it is that the two samples were drawn from the same distribution.

Figure	**D-statistic**	**P-value**
Men-Women	0.375	0.0678
SA Men-Women	0.166	0.9024
Japan Men-Women	0.541	0.0014

Table 4: KS-test values for men-women circadian rhythms

From table 4, we observe that the p-value for Men-Women is greater than the significant value of 0.05. This means that the circadian rhythms for these two groups are similar and we are able to reject the null hypothesis that these are drawn from the same distribution. We owe this difference to the gender roles associated with women to a varying degree in all countries. It is reported that gender does not affect sleep cycle in Blacks [8] which is in line with our KS-test values. Japan ranks high in Gender Inequality Index which could explain the less sleep time that women enjoy compared to men.

To test the rhythmicity of circadian rhythms worldwide, we sampled the date range in our dataset for a period of 7 days and calculated the average twitter activity for a week. Figure 6 demonstrates that the series is repeating in structure ev-

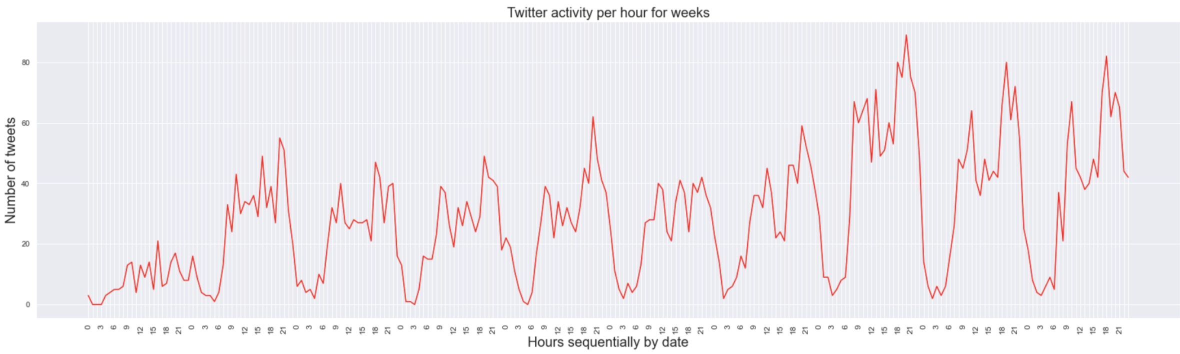

Figure 17: Twitter activity for all users

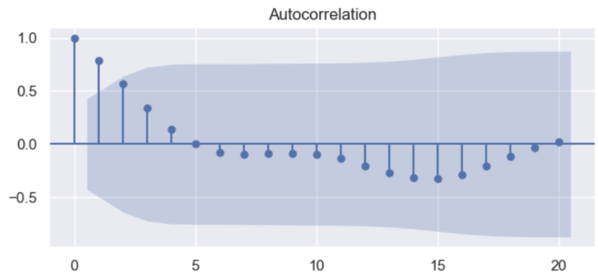

Figure 18: Autocorrelation to determine cycle length

eryday with higher activity on some days. This is further validated by the autocorrelation visualization, figure 18. Since, there is a periodical event that impacts current data heavier then the rest ACF seems to be the best measure of periodicity detection. The first point (with index 0) has a height 1 as current value always explain fully current value. It is evident from the plot that there is a periodicity which is repeated around the 14th day.

7 Bias Mitigation

We studied only the top most populous countries in our dataset since we do not want to generalize hypotheses around countries that have a sparse user base. Also, it is not possible to study the demographics of 200 countries. Skewed distribution due to less representation in the dataset will be an issue.

While tagging locations several human level biases seem to have occurred. Some locations were written in Hindi and the annotators being Indians were able to recognise them. Places tagged in languages that annotators could not identify were not tagged. Further, some users put celebrity profile images,model predicts incorrect demographics. This lead to bias in the inferred age and gender. Some post images of their younger selves. Age prediction in such cases might not be the correct. To prevent bias in timestamps, we first calculated the time in UTC and to their local time zones. This allows us to scale all tweets to 24 hour noon-to-noon cycle. Initially, we analysed two weeks worth of data for any user, which caused discrepancies since users had a non-uniform distribution of tweets across 14 days, meaning, some users had 10 tweets in 2 weeks while others had 200+. To normalize such sparsity, we collected 140 tweets from each user. This also prevents bias due to nation specific Twitter usage, i.e., two weeks in different countries will have different volumes of activity based on what time of the year or month it is. For instance, the end of the 14 days period for which we scraped the data will see a rise of tweets across the USA because of Halloween. However, since we are collecting 140 latest tweets from each of our users, the timeline does not matter.

8 Conclusion

Our current work shows a range of Twitter usage analyses across various socio-demographic elements how we can benefit from such data to determine the sleep cycle of individuals. We create our own dataset with age, gender, nationality and twitter usage of individuals and plot their circadian rhythms as determined by twitter activity. We further justify the strength of the rhythms using KS-test and Autocorrelation Function. With every analysis we refer existing research to justify the claim.

As a future work, we would like to study why there is a periodicity of 14 days in Twitter Sleep Cycle. We would also like to statistically distinguish between night owls and outliers.

References

[1] (2008)kolmogorov–smirnov test. in: The concise encyclopedia of statistics. springer, newyork,ny.,https://doi.org/10.1007/978-0-387-32833-1$_2$14. 2008.

[2] Hiwot Berhanu, Andualem Mossie, Samuel Tadesse, and Daniel Geleta. Prevalence and associated factors of sleep quality among adults in jimma town, southwest ethiopia: a community-based cross-sectional study. *Sleep disorders*, 2018, 2018.

[3] William J Chopik. The benefits of social technology use among older adults are mediated by reduced loneliness. *Cyberpsychology, Behavior, and Social Networking*, 19(9):551–556, 2016.

[4] Brenda Curtis, Salvatore Giorgi, Anneke EK Buffone, Lyle H Ungar, Robert D Ashford, Jessie Hemmons, Dan Summers, Casey Hamilton, and H Andrew Schwartz. Can twitter be used to predict county excessive alcohol consumption rates? *PloS one*, 13(4):e0194290, 2018.

[5] Johannes C Eichstaedt, Hansen Andrew Schwartz, Margaret L Kern, Gregory Park, Darwin R Labarthe, Raina M Merchant, Sneha Jha, Megha Agrawal, Lukasz A Dziurzynski, Maarten Sap, et al. Psychological language on twitter predicts county-level heart disease mortality. *Psychological science*, 26(2):159–169, 2015.

[6] Salvatore Giorgi, Sharath Chandra Guntuku, Johannes C Eichstaedt, Claire Pajot, H Andrew Schwartz, and Lyle H Ungar. Well-being depends on social comparison: Hierarchical models of twitter language suggest that richer neighbors make you less happy. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 15, pages 1069–1074, 2021.

[7] Leila Hedayatifar, Rachel A Rigg, Yaneer Bar-Yam, and Alfredo J Morales. Us social fragmentation at multiple scales. *Journal of The Royal Society Interface*, 16(159):20190509, 2019.

[8] Dayna A Johnson, Chandra L Jackson, Natasha J Williams, and Carmela Alcántara. Are sleep patterns influenced by race/ethnicity–a marker of relative advantage or disadvantage? evidence to date. *Nature and science of sleep*, 11:79, 2019.

[9] Jason J Jones, Gregory W Kirschen, Sindhuja Kancharla, and Lauren Hale. Association between late-night tweeting and next-day game performance among professional basketball players. *Sleep Health*, 5(1):68–71, 2019.

[10] Hiroyuki Kikuchi, Yuko Odagiri, Yumiko Ohya, Yutaka Nakanishi, Teruichi Shimomitsu, Töres Theorell, and Shigeru Inoue. Association of overtime work hours with various stress responses in 59,021 japanese workers: retrospective cross-sectional study. *PloS one*, 15(3):e0229506, 2020.

[11] Yue Leng, Nick WJ Wainwright, Francesco P Cappuccio, Paul G Surtees, Robert Luben, Nick Wareham, Carol Brayne, and Kay-Tee Khaw. Self-reported sleep patterns in a british population cohort. *Sleep medicine*, 15(3):295–302, 2014.

[12] Yue Leng, Nick WJ Wainwright, Francesco P Cappuccio, Paul G Surtees, Robert Luben, Nick Wareham, Carol Brayne, and Kay-Tee Khaw. Self-reported sleep patterns in a british population cohort. *Sleep medicine*, 15(3):295–302, 2014.

[13] Eugene Leypunskiy, Emre Kıcıman, Mili Shah, Olivia J Walch, Andrey Rzhetsky, Aaron R Dinner, and Michael J Rust. Geographically resolved rhythms in twitter use reveal social pressures on daily activity patterns. *Current Biology*, 28(23):3763–3775, 2018.

[14] Monica P Mallampalli and Christine L Carter. Exploring sex and gender differences in sleep health: a society for women's health research report. *Journal of women's health*, 23(7):553–562, 2014.

[15] Karl Peltzer. Differences in sleep duration among four different population groups of older adults in south africa. *International journal of environmental research and public health*, 14(5):502, 2017.