# Wikipedia Watching

We are interested in keeping tab on the changes being made to Wikipedia. Wikipedia provides a Streaming API using which we can see the changes being made to various Wikipedia sites in realtime. We will use it to generate various reports.

## Task 1

We will start small. To begin with, let us keep track of the data coming on the API and every 1 minute, generate the two reports described below. The reports should be based on the data received in the last 1 minute. The reports should be printed on the terminal/console.

## Task 2

While reports based on 1 minute data are useful, we would like to keep track of the changes over a larger window of time. We will now start generating reports based on 5 minutes worth of data instead of 1 minute. The reports will still be generated every minute.

Ex:
Minute 1 Report - Minute 0-1 data
Minute 2 Report - Minute 0-2 data
…
Minute 5 report - Minute 0-5 data
**Minute 6 report - Minute 1-6 data (**take note**)**
Minute 7 report - Minute 2-7 data
…

Every 1 minute, print the same reports as before but now the reports should be based on the data received in the last 5 minute.

## Reports

### Domains Report

Print the number of the Wikipedia domains that have been updated, followed by a list of the domains sorted by the count of how many unique pages were updated on each. Pages with the same title are assumed to be the same.

### Sample Report:

```
Total number of Wikipedia Domains Updated: 10

en.wikipedia.org: 10 pages updated
es.wikipedia.org: 6 pages updated
ru.wikipedia.org: 4 pages updated
hi.wikipedia.org: 1 page updated
...
```

## Users Report

Print a list of users that have made changes to **en.wikipedia.org** domain, sorted by their total edit count (available as performer->user_edit_count in each event). If the same user shows up multiple times in the given time period, then use the highest edit count seen for them.

Apart from regular users, various bots also make changes to Wikipedia pages. **For generating this report, any bot users should be excluded.** Whether a user is a bot or not is mentioned in the API response.

### Sample Report:

```
Users who made changes to en.wikipedia.org
James123: 12012
Jiten_Sharma: 8121
...
```

# Wikipedia Event Stream API

Use the data provided by the https://stream.wikimedia.org/v2/stream/revision-create endpoint. This provides a real time feed of all the new revisions being created on Wikipedia.

To get started, you can find Javascript and Python example code and link to libraries here:
https://wikitech.wikimedia.org/wiki/Event_Platform/EventStreams

You can see the type of data you will receive from the API here:
https://stream.wikimedia.org/v2/ui/#/?streams=revision-create

# Notes

- You can complete the task in the programming language of your choice.

# Submission

Setup a local Git repo with your code. The repo should include a README file with instructions on how to setup and run the code. Make a compressed .zip or .tar.gz archive of the repo and send it as an email attachment.