Step 1: We used TripAdvisor data from

<https://www.tripadvisor.com/Hotel_Review-g60745-d6485213-Reviews-The_Verb_Hotel-Boston_Massachusetts.html>

Step 2: Created fake reviews as follows:



Step 3: Scraped Reviews using following automation

<http://localhost:8888/notebooks/Downloads/ADM_Assignment2.ipynb>

Step 3: Set up MTurk job for a user to vote if the review is fake or real.

 

Step 4: Created a Tableau worksheet stating real and fake reviews.

<https://public.tableau.com/authoring/Book2_16367846000990/Sheet1#1>

Graphical user interface

Description automatically generated with medium confidence

Step 5: Automated MTurk job

Part 2:

Step 5: Created a S3 bucket and uploaded the mturk results on the bucket

<http://localhost:8888/notebooks/Downloads/boto3.ipynb>

Step 6: Created rule-based engine filter fake and real reviews based on following rules

<http://localhost:8888/notebooks/Downloads/ADMAssignment2_Part2.ipynb>

1. Length of review >100 characters
2. Created a list of offensive words to check for fake reviews against the mturk results.



1. Length of review <500 characters
2. Check for URLs in the review. If the review has URLs, flag else publish it.

Step 6: Merge all the file from 9 groups and run the file on the rule-based engine and show the findings.

