**Data, Evaluation, and Results**

[**Results**](https://docs.google.com/spreadsheets/d/11k0XKaYwJ-Xh-C9sm5M_a6kfQJCsK0PlPTLJs1dDZrU/edit#gid=0)

**Data:**

* [Training and test data with stance gold labels and additional labels](http://alt.qcri.org/semeval2016/task6/data/uploads/stancedataset.zip) such as 'target of opinion' and 'sentiment' . These additional annotations were not part of the SemEval-2016 competiton, but are made available for future research. Details about this dataset are available in this paper:  
    
  Saif M. Mohammad, Parinaz Sobhani, and Svetlana Kiritchenko. 2016. [Stance and sentiment in tweets.](http://www.saifmohammad.com/WebDocs/1605.01655v1.pdf) Special Section of the ACM Transactions on Internet Technology on Argumentation in Social Media, 2017, 17(3).
* [Test data with gold labels](http://alt.qcri.org/semeval2016/task6/data/uploads/semeval2016-task6-testdata-gold.zip)
* [TEST DATA](http://alt.qcri.org/semeval2016/task6/data/uploads/semeval2016-task6-testdata.zip) without gold labels (Submissions are due **Jan. 18, 2016, 11:59PM Pacific Standard Time (GMT-8)**)
* [Trial data](http://alt.qcri.org/semeval2016/task6/data/uploads/semeval2016-task6-trialdata.txt) ([readme](http://alt.qcri.org/semeval2016/task6/data/uploads/semeval2016-task6-trialdata-readme.txt)) - can be used for training
* [Training data for Task A](http://alt.qcri.org/semeval2016/task6/data/uploads/semeval2016-task6-trainingdata.txt)  ([readme](http://alt.qcri.org/semeval2016/task6/data/uploads/semeval2016-task6-trainingdata-readme.txt))
* [Domain corpus for Task B](http://alt.qcri.org/semeval2016/task6/data/uploads/semeval2016-task6-domaincorpus.zip) (note: it takes about 5 days to download the tweets in the corpus)

**Data Visualization:**

* An [Interactive Visualization](http://www.saifmohammad.com/WebPages/StanceDataset.htm) of the Stance Datset is now available. It shows various statistics about the data.
* Note that it also shows sentiment and target of opinion annotations (in addition to stance).
* Clicking on various visualization elements filters the data. For example, clicking on 'Feminism' and 'Favor' will filter all sub-visualizations to show information pertaining to only those tweet that express favor towards feminism. You can also use the check boxes on the left to view only test or training data, or data on particular targets.

**Instructions to Annotators**

* We used this [questionnaire](http://alt.qcri.org/semeval2016/task6/data/uploads/stance-question.pdf) to obtain annotations
* Annotators were restricted to those living in USA
* For the classification tasks A and B, options 3 and 4 listed in the questionaire were collapsed into one class NONE (neither favor nor against)

**Evaluation:**

* We will use the macro-average of F-score(FAVOR) and F-score(AGAINST) as the bottom-line evaluation metric.
* [Evaluation Script v2](http://alt.qcri.org/semeval2016/task6/data/uploads/eval_semeval16_task6_v2.zip)  (last updated: January 11th, 2016) (the same script can be used for both task A and Task B)  
    You can use it to:  
    -- check the format of your submission file  
    -- determine performance when gold labels are available (note that you can also use the script to determine performance on a held out portion of the training data to gauge your system's progress)
* [A separate evaluation script](http://alt.qcri.org/semeval2016/task6/data/uploads/evaluation-code-for-scores-on-different-subsets-of-test-set.zip) to determine scores on the following subsets of the test set: (a) where the given target of interest is the same as the target of opinion in the tweet, and (b) where the given target of interest is \*not\* the target of opinion in the tweet. NOTE: This script was not part of the official competition, and is provided only as a means for further analysis of the results.