

The Effects of Course Variabilities on Students' Sentiments

BY SAMER AL MASRI

Contributions

Analysis of the effects of variabilities on the overall sentiment.

- Proving and disproving hypothesis about the data and checking their generality among other courses.

Testing the IBM AlchemyLanguage Tone Analyzer's accuracy.

Publishing the code used to analyze this data on Github.



IBM Watson Tone Analyzer

Emotion

Joy

Fear

Sadness

Disgust

Anger

Language Tone

Analytical

Confidence

Tentative

Social Tone (big 5 personality traits)

Openness

Conscientiousness

- acting in an organized way

Extraversion

- being an extrovert

Agreeableness

- cooperativeness

Emotional Range or Neuroticism

- emotional sensitivity

Source: <https://www.ibm.com/watson/developercloud/doc/tone-analyzer/understand-tone.html>

Information About Courses

1. Introduction to Software Product Management (3 weeks)
Consists of: exams, supplements, lectures
2. Software Processes and Agile Practices (5 weeks)
Consists of: exams, supplements, lectures
3. Client Needs and Software Requirements (5 weeks)
Consists of: quiz, exam, **peer work**, supplement, lecture, others
4. Agile Planning for Software Products (5 weeks)
Consists of: exam, supplement, **peer work**, lecture
5. Reviews and Metrics for Software Improvements (5 weeks)
Consists of: exams, supplements, lectures
6. Software Product Management Capstone (7 weeks)
Consists of: exam, **peer work**, supplement, lecture

Top Countries Where Students are Living

	Introduction to Software Product Management	Software Processes and Agile Practices	Client Needs and Software Requirements	Agile Planning for Software Products	Reviews and Metrics for Software Improvements	Software Product Management Capstone
Country #1	United States	United States	United States	United States	United States	United States
Country #2	India	India	India	India	India	United Kingdom
Country #3	Ukraine	Ukraine	Ukraine	Ukraine	Ukraine	Germany
Country #4	Russian Federation	Russian Federation	Russian Federation	Russian Federation	Russian Federation	Russian Federation
Country #5	Brazil	Brazil	Brazil	Germany	Germany	Netherlands
Country #6	United Kingdom	Spain	United Kingdom	Spain	Brazil	Bulgaria

Course Population

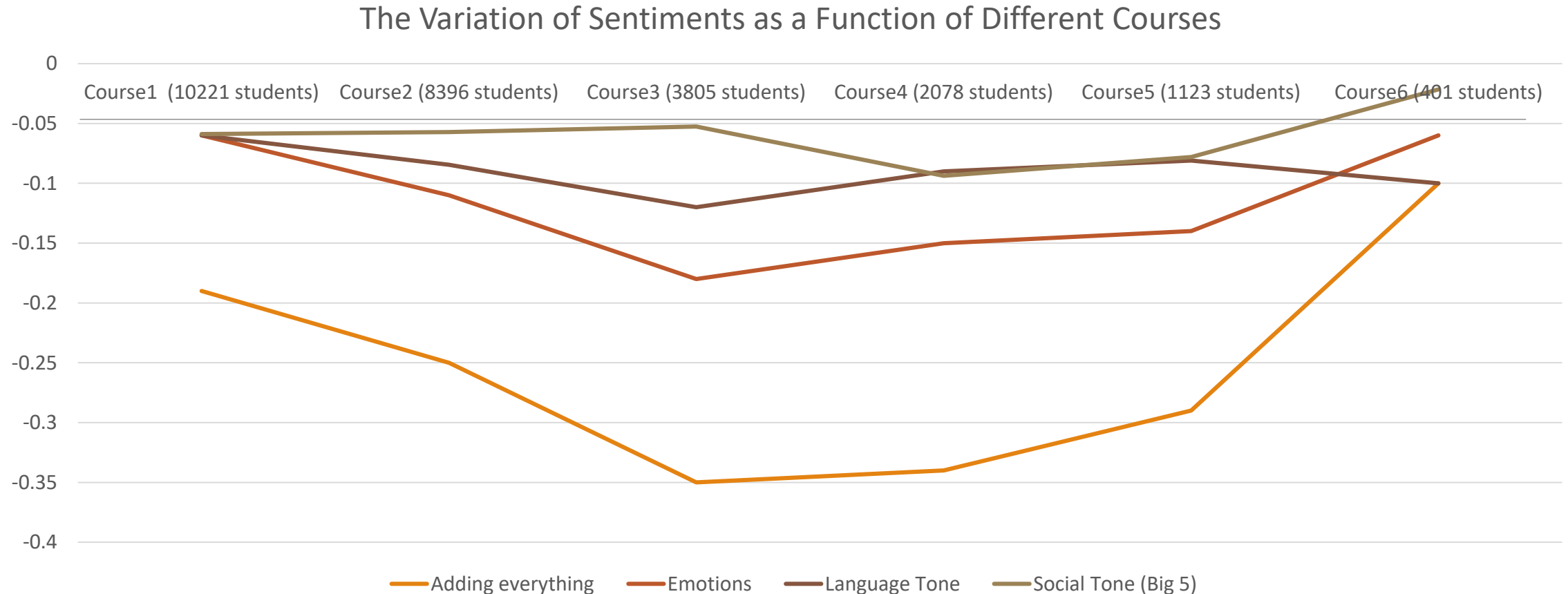
Courses	Hispanic Origin %	Education Level	Gender	Age
Course 1	13.33%	MSc – BS	75.09% males	35.0178
Course 2	16.06%	MSc – BS	75.25% males	35.24
Course 3	15.64%	MSc – BS	72.99% males	35.405
Course 4	14.51%	MSc – BS	73.11% males	35.16
Course 5	14.48%	MSc – BS	72.78% males	35.74
Course 6	20.37%	MSc – BS	75.92% males	36.88

- Age is increasing as the courses advance.
- Course 6 has more males and more Hispanic Origin people than Course 1.

Different Criteria for Sentiment Analysis

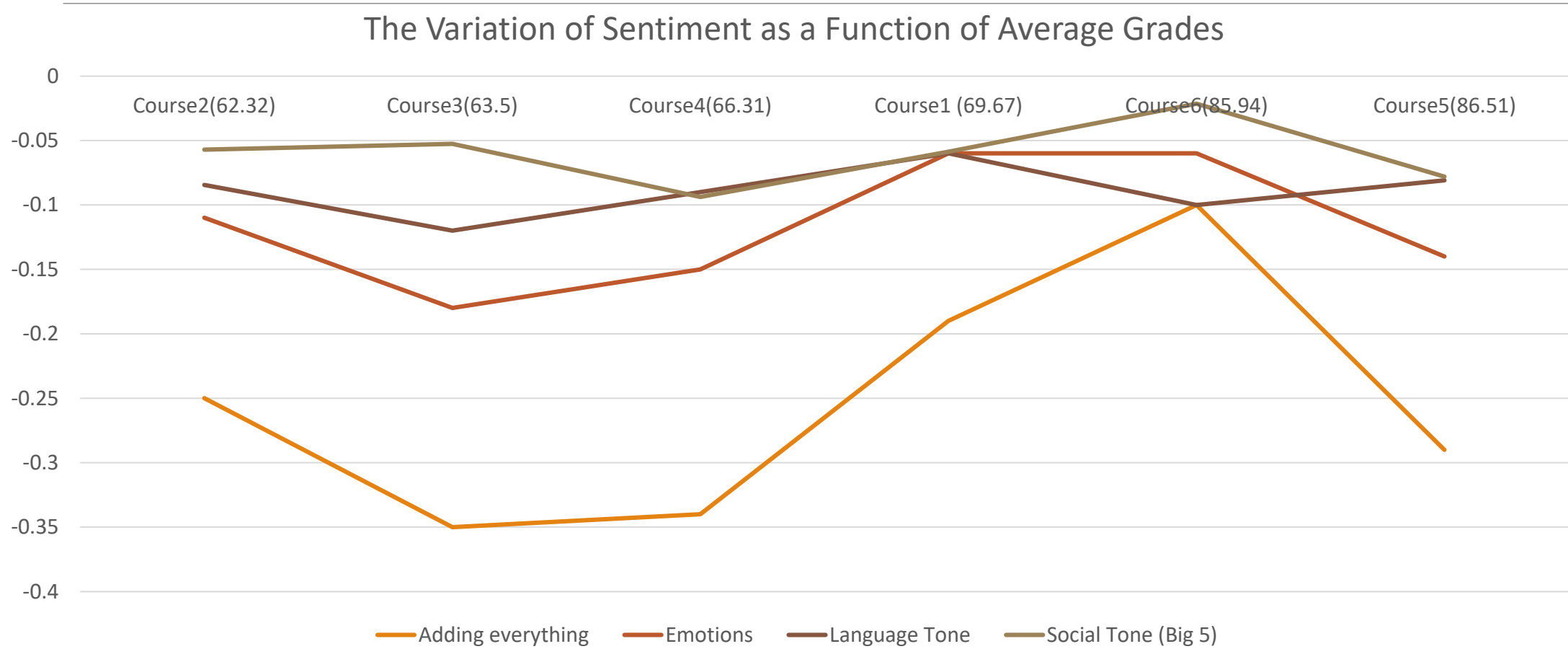
Course Number	Adding Everything	Emotions	Language Tone	Social Tone (Big 5 Personality Traits)
Course 1	-0.19	-0.06	-0.06	-0.0588
Course 2	-0.25	-0.11	-0.0845	-0.0571
Course 3	-0.35	-0.18	-0.12	-0.0526
Course 4	-0.34	-0.15	-0.09	-0.0938
Course 5	-0.29	-0.14	-0.809	-0.0781
Course 6	-0.10	-0.06	-0.10	-0.0214

Visual Representation of the Sentimental Analysis

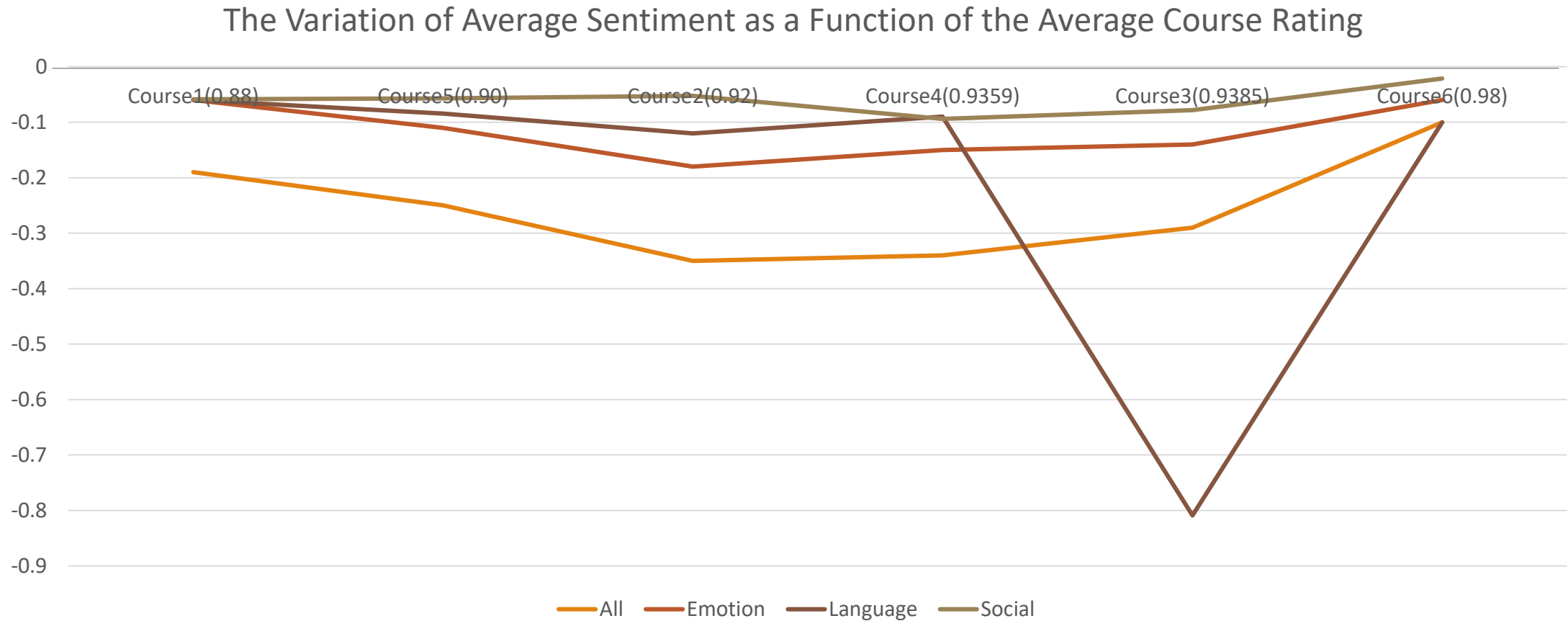


- As student number decreases, social tone is increasing.
- Having peer work decreases the language tone (Courses 3, 4, and 6).

Effect of grades on Sentiment

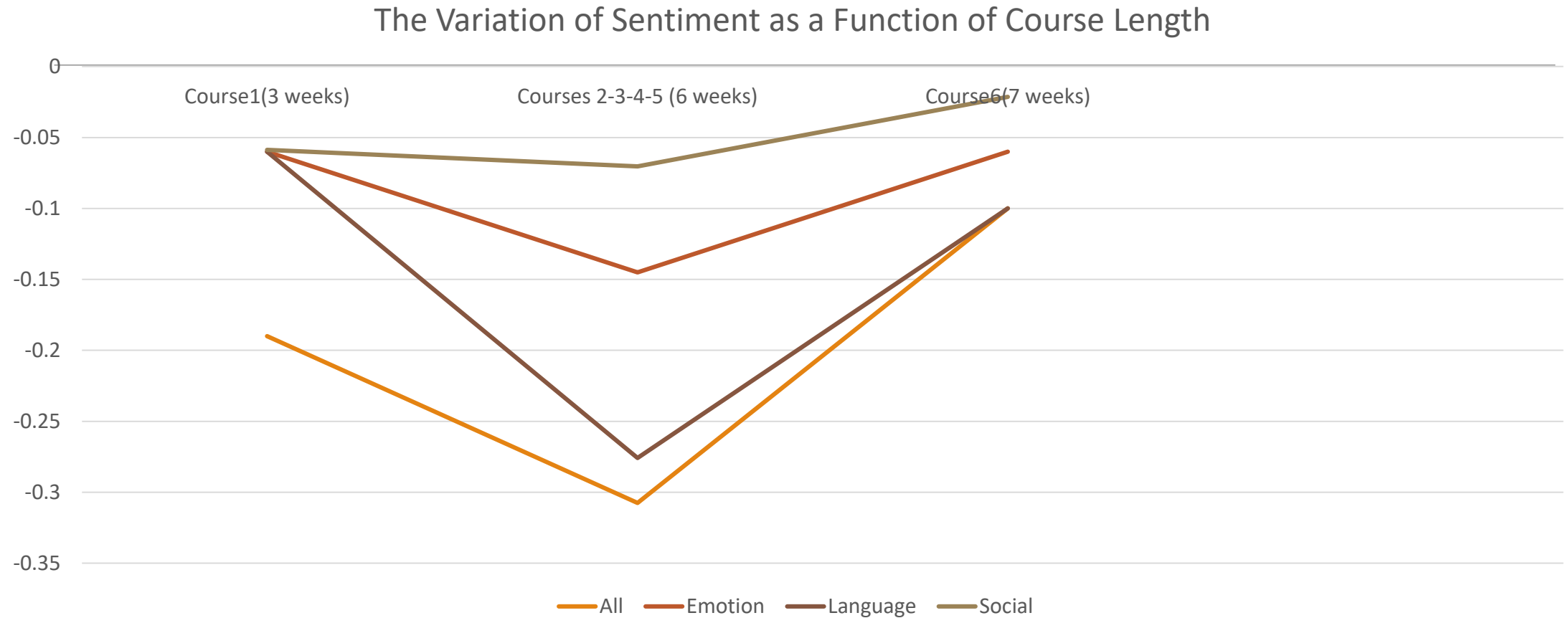


Correlation Between Sentiment and Rating of Course



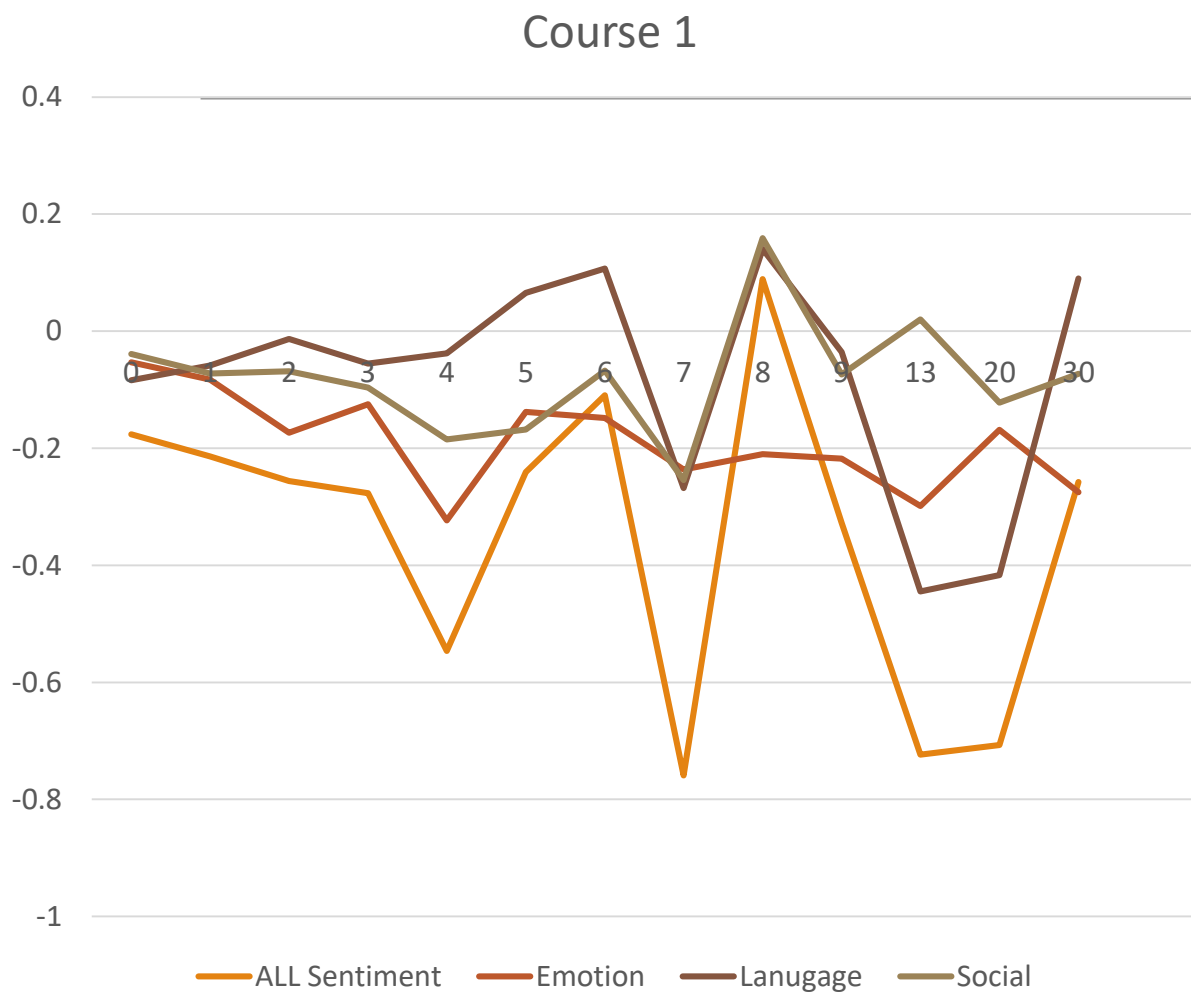
- Students' course rating isn't directly impacted by their feeling about the course.

The Effect of Course Length on Sentiment



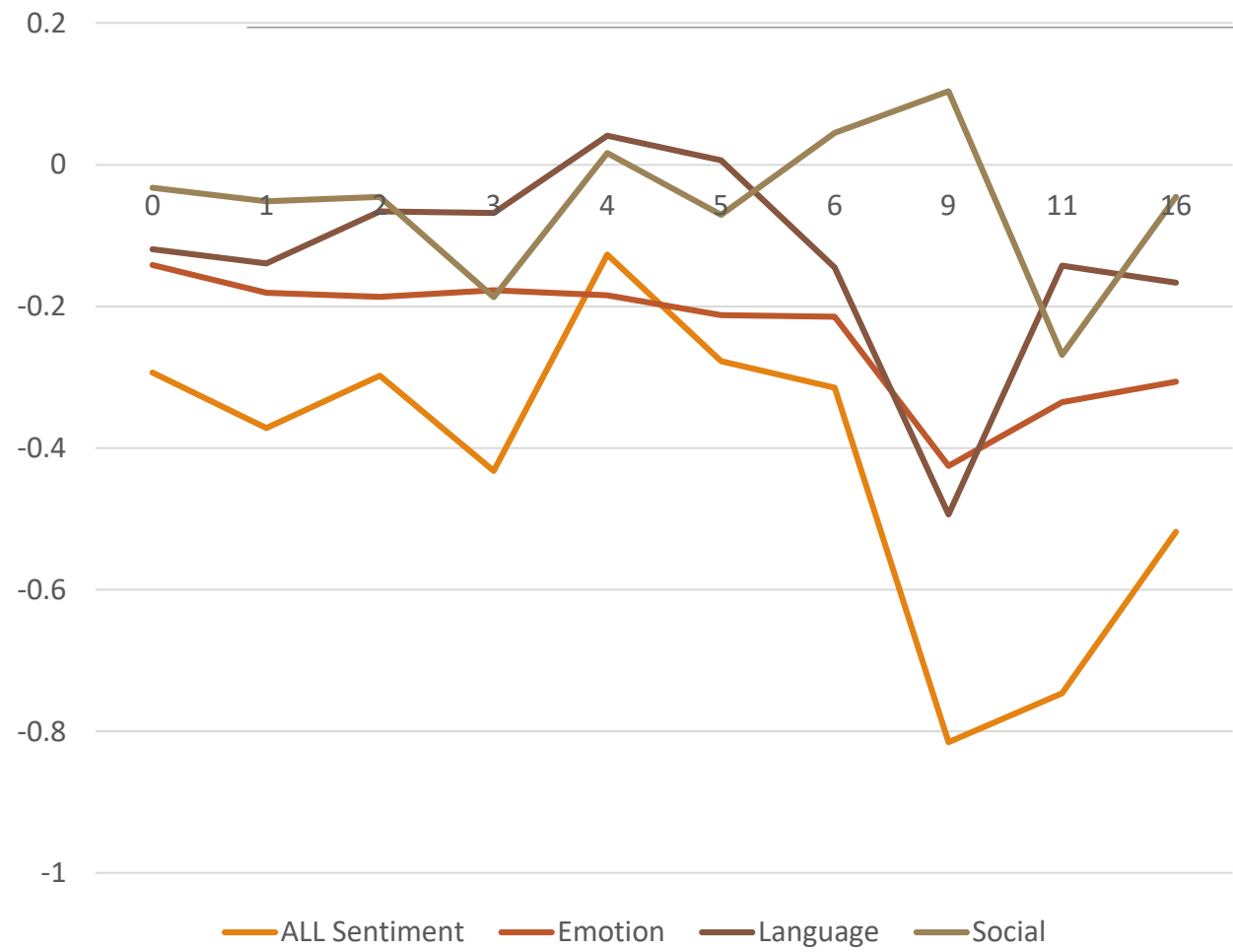
- The sentiment decreases when the course duration increases.

Variation of Sentiment as a Function of Post Ratings

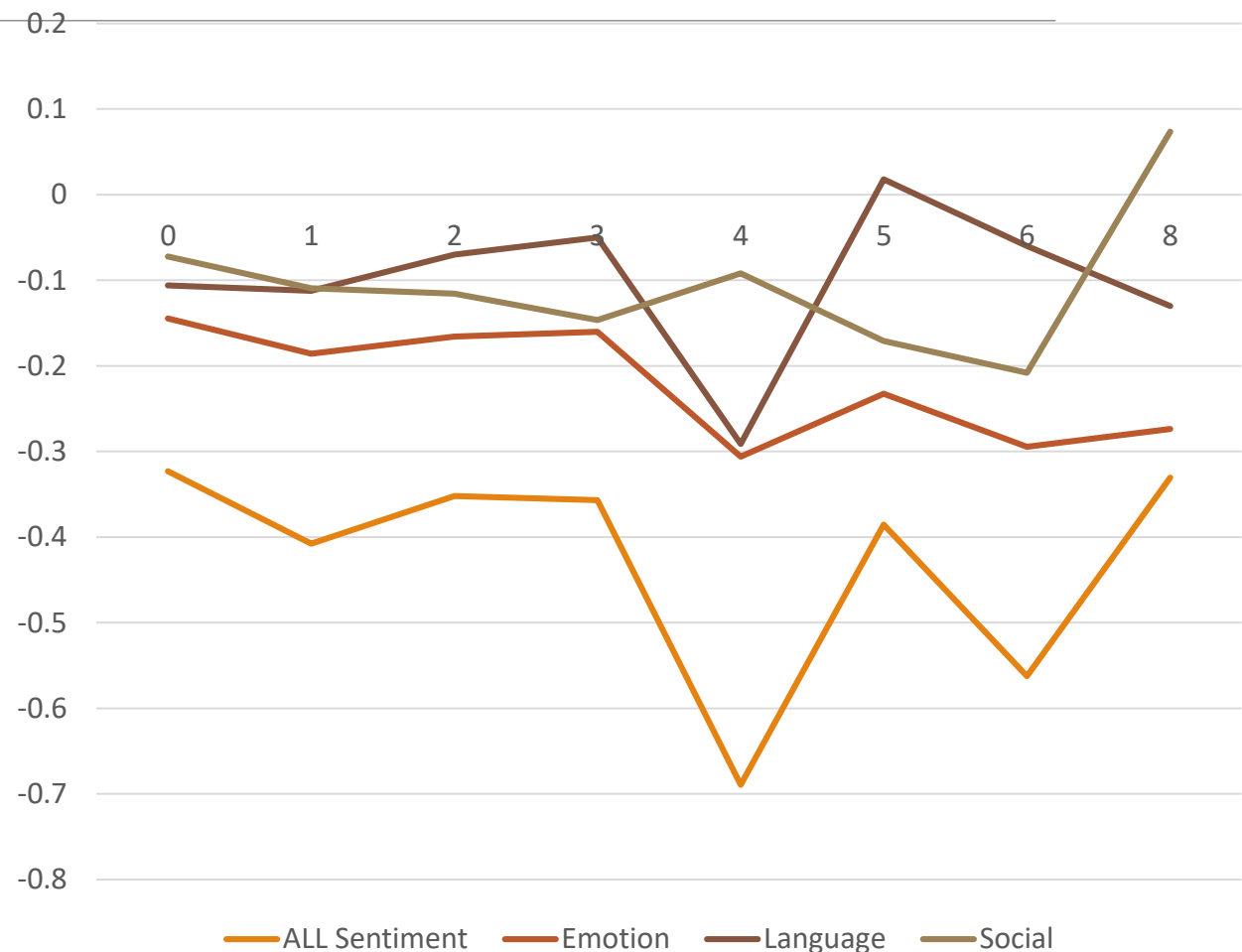


Variation of Sentiment as a Function of Post Ratings (Continued)

Course 3

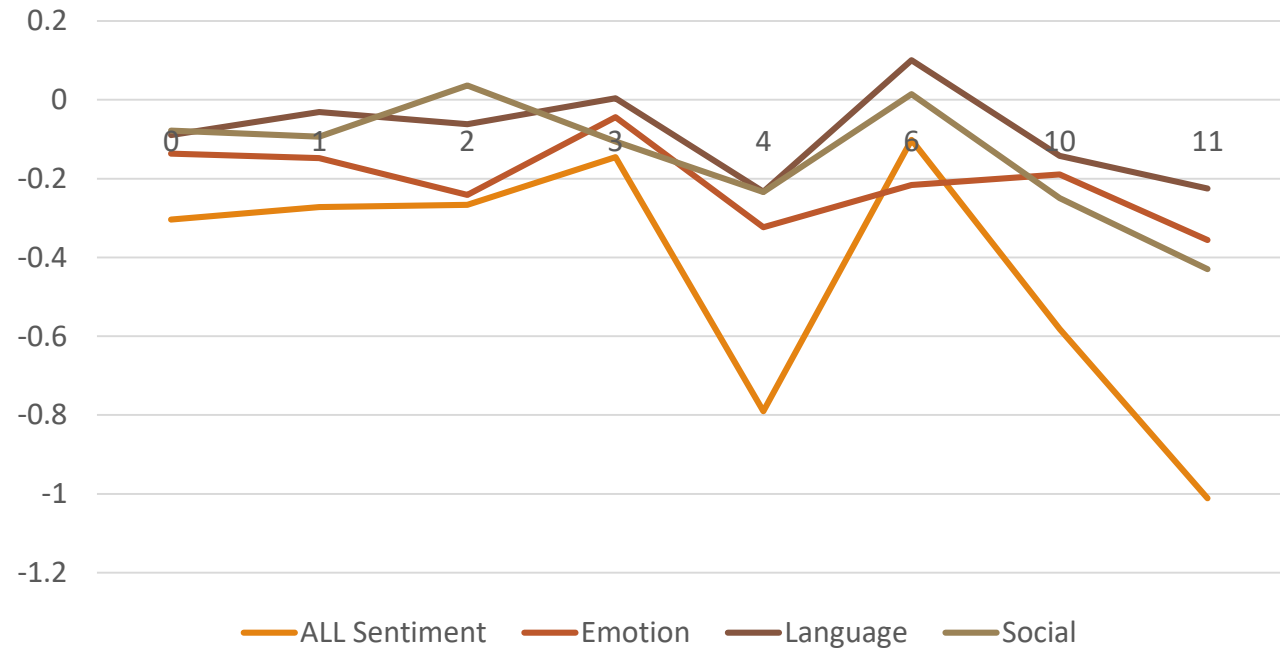


Course 4

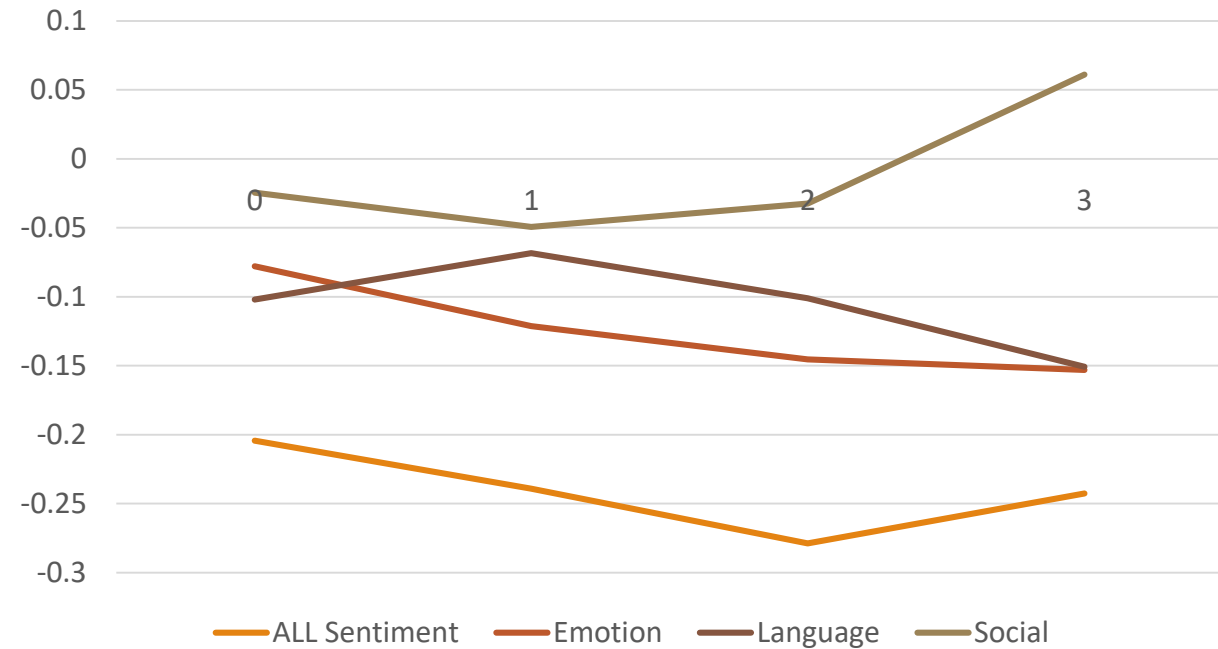


Variation of Sentiment as a Function of Post Ratings

The Variation of Sentiment as a Function of Post Ratings



The Variation of Sentiment as a Function of Post Ratings



- The peer work is negatively affecting the language sentiment.

Challenges

Deciding how to calculate the average Tone.

Student IDs are not connectable throughout different tables.

Technical advice: It is easier to use python for processing CSV and JSON files.

The Guide document included didn't explain about peer work (graded peer, phased peer...).

Some discussion answer IDs are found in the discussion_answer_votes.csv but not in discussion_answers.csv.

Future Work

Manually checking posts to have an idea how accurate the IBM sentimental analysis is.

Outsourcing the post checking to get more accurate results.

The code is tailored to analyze any course data from Coursera, hence more courses can be analyzed to get more accurate hypothesis.