Good morning, distinguished professors. I am Zhaoyang Tian. My paper’s topic is Emotional Analysis & Evaluation of Pros and Cons Regarding Students’ Activities ---- Based on Comprehensive Education Evaluation System in Tsinghua University High School.

This is the agenda of our presentation. I will mainly discuss the following eight sections.

To begin with, I will introduce the research background, purpose, and general process. Sponsored by Beijing Education Committee, Tsinghua University High School has developed Comprehensive Quality Evaluation Platform for Beijing General High School Students, which aims that students can voluntarily upload the photos and text description and reveal their special activities and strength that cannot be reflected in class or by scores. The CQE system has been implemented for over a year, while no research towards the data in CQE has been made.

The past research mostly focus on methodology and unique example conducted by a few researchers, with the limitation of small amount of data and specific experiment. In light of the condition, this paper conducts research in the hope of offering a practical solution in a statistical way.

This paper demonstrates the data set in a brand new manner, evaluates the positive and negative feelings behind the data, and recommends the activities for the school to refer to.

This is the general process of our research. After sales records from our school, I first segment the sentences into words. Then we continue processing the data and analyze the frequency of the words to have a rough impression. Next, we assemble some wordclouds according to the frequency to demonstrate the data in a new way. After that, we create an original method to analyze the positive and negative feelings, and uses SnowNLP to bolster the method. Finally, we apply support vector machine to extract the words that are critical to determine the activities that can bring intense feelings to students. Finally, I manually search for the positive and negative aspects of the popular activities.

I obtains the data in the CQE system of grade G17 in Tsinghua University High School from February, 2018 to June, 2018. Each pieces of data is an activity uploaded by the students. This paper has 6 dimensions to study.

This paper apply JIEBA segmentation and word frequency analysis to the data and find out the most welcomed activities. With this method, this paper finds the activities that are mostly mentioned, which represents welcomed activities. This paper maps the data into a bar chart to demonstrate the data. WE can see that The sports games, basketball match, and the soccer match are the most frequent activities.

With the aid of WORDCLOUD module, this paper generates wordclouds through the set of parameters and load the data based on the frequency calculated above.

This paper has also created an original method to evaluate the emotional quantity of sentences. First, this paper enumerates a positive dictionary and a negative dictionary. The positive emotional words are such: fine, good, magnificent, gorgeous, so and so forth; the negative emotional words such as: bad, ill, wicked, so and so forth. This paper devises a set of methods to define positive and negative. If there is a positive word, the value of positive score will add 1. If there is a negative word, the value of negative score will add 1. This paper finds the degree words before the emotional words. Then this paper sets weights for degree words and multiply them by emotional values. If an exclamation mark is found in a sentence, the positive value will be added by 2. This paper finds the reverse words before the emotional words. If the number of the reverse words is odd, the score of the word is multiplied by - 1. If it is even, the score of the word will remain unchanged.

This paper examines the correlation coefficients between the score of SnowNLP and the positive score of Original Emotion Evaluation Method to bolster the proposed method.

Since there are not a large number of it, which is about 10-20 pieces of data each, this paper decides to find the positive and negative aspects of the activities manually.

“Nervous”(紧张) is a word that is often mentioned in pieces of data that are more likely to be treated as negative in sports activities, which includes sports meeting, soccer match, and basketball match.

If there are too less words in a piece of data, it will tend to be viewed as a negative one, because it contains too little information which means the hosts of the data may be less involved in the activity. For instance, if a person simply participates in the admission of a sports meeting or merely works as a staff, it is probable that the data is viewed as negative.

If a person describe an experience as something like “visit”(参观), “observe”(观察), it will be more likely to be viewed as negative than that described as something like “participate”(参加), “involvement”(体验).

Data related to innovative achievements, excursion experience, club events are very likely to be viewed as positive.

A piece of descriptive data is more likely to be positive than a piece of narrative data.