# Text Analysis of Data

Abstract: This paper delves into the high-frequency keywords of Tang and Song poetry using Python programming language. Through the analysis and comparison of the top ten keywords in frequency, we gain insight into the unique characteristics of each era. Tang poetry is known for its celebration of nature, with keywords such as "white clouds," "spring breeze," and "bright moon" painting a vivid picture of the beauty of the natural world. On the other hand, Song ci, a type of poetry that emerged during the Song dynasty, often conveys a sense of melancholy and sadness, as seen in the high frequency of keywords such as "plum blossom" and "west wind." Furthermore, the presence of keywords such as "fengliu" and "xiangsi" in Song ci reflects the cultural values and emotional experiences of the Song dynasty society. "Fengliu" embodies the pursuit of elegance and romance, while "xiangsi" expresses the poet's yearning and pursuit of love. These keywords provide a glimpse into the emotional landscape of the Song dynasty. Through our analysis, we gain a deeper understanding of the cultural and historical context of Tang and Song poetry. Moreover, the findings of this research provide valuable insights for further exploration of Chinese literary history. By using computational tools to analyze literary works, we can uncover hidden patterns and trends that may not be immediately apparent through traditional methods. This study demonstrates the potential of digital humanities research in enhancing our understanding of literature and culture. Ultimately, this paper sheds light on the power of technology to illuminate the nuances and complexities of artistic expression.

Introduction: Tang and Song poetry are two important periods in the history of Chinese literature. They not only produced a large number of outstanding poets and lyricists but also left many classic poetry works. These poems are not only the essence of literary art but also an important part of Chinese culture. However, due to the huge differences in social background, cultural atmosphere, and aesthetic orientation between the Tang and Song dynasties, the cultural characteristics and artistic styles of Tang and Song poetry are also very different. This paper aims to explore the cultural characteristics and historical background of Tang and Song poetry through comparative studies of high-frequency keywords, in order to better understand the development and evolution of Chinese literary history. To this end, we have written a program using the Python language to statistically analyze the high-frequency keywords in Tang and Song poetry. By comparing the top ten keywords in frequency in Tang and Song poetry, we found significant differences in the high-frequency keywords between the Tang and Song dynasties. The results of this study not only help us better understand the cultural characteristics and historical background of Tang and Song poetry but also provide useful references for us to further explore Chinese literary history.

	唐朝关键词	词频	宋朝关键词	词频
1	今日	1156	东风	1286
2	春风	1126	何处	1161
3	白云	1080	人间	1074
4	千里	1049	风流	828
5	长安	916	相思	749
6	故人	882	梅花	755
7	人间	867	春风	746

| 8 | 无人 | 849 | 西风 | 746 |
| 9 | 明月 | 808 | 归来 | 745 |
| 10 | 不得 | 808 | 归去 | 689 |

唐朝关键词柱状图

宋朝关键词柱状图
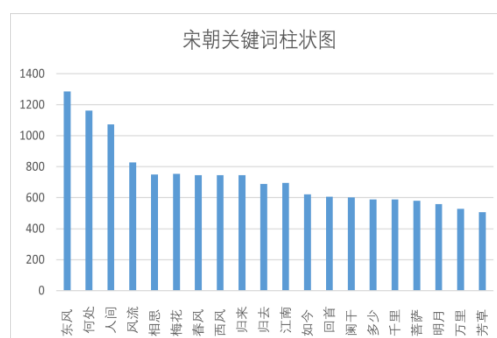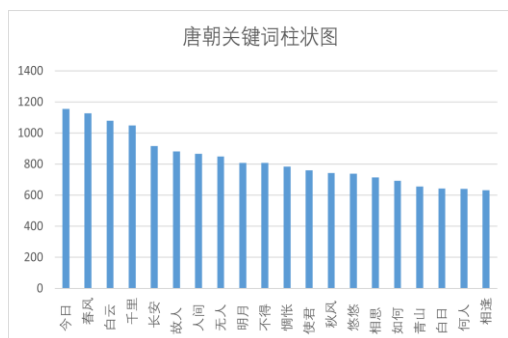
From Table 1, it can be seen that the data presented through calculations can provide some clues for humanistic researchers, inspiring many questions, such as: among the top ten highest frequency words in Tang and Song poetry, only "spring breeze" and "human world" overlap. What causes these two words to appear so frequently in two eras with vastly different social backgrounds and cultural atmospheres? What are the specific similarities and differences in their meanings and contexts when they appear in poetry in the two eras? What differences do they reflect between Tang and Song? These are all questions worth exploring. For example, in Tang poetry, does "today" reflect the high attention of the Tang people to the present and reality, and is it a manifestation of the openness and prosperity of the Tang dynasty? After all, people tend to focus on real life when they are in a prosperous era, and they can place "today" in an important position to praise and sing. At the same time, the high frequency appearance of "Chang'an" is very striking. As the capital of the Tang dynasty, Chang'an was the center of China's politics, economy, and culture at that time, and one of the most prosperous and culturally rich cities in the world, naturally becoming the place that literati and poets dreamt of. Tang poets mentioned "Chang'an" extensively in their poetry, which on the one hand reflected their infinite yearning and admiration for the capital city, and on the other hand directly reflected the prosperity of Chang'an itself. This is undoubtedly a powerful example of the atmosphere of the flourishing Tang dynasty.

In addition, "white clouds," "spring breeze," and "bright moon" are all natural landscapes that often carry a certain positive meaning when used, such as praising the beauty of spring and clear scenery and expressing a feeling of relaxation and comfort. These two keywords occupy a prominent position among the top ten in terms of frequency, which is enough to show the Tang poets' love and attention to natural scenery. In contrast, in Song ci, natural landscape and object keywords with negative connotations such as "plum blossom" and "west wind" appear in the top ten frequently used keywords, which are often used to express a sense of desolation and sadness, opposite to the positive atmosphere shown in Tang poetry. "Fengliu" (elegance and romance) and "xiangsi" (longing and yearning) frequently appear in Song ci, reflecting to some extent the cultural characteristics and emotional experiences of the Song dynasty society. "Fengliu" embodies the pursuit of elegance and romance by the literati and scholars of the Song dynasty, while "xiangsi" expresses the poet's yearning and pursuit of love. Both of these reflect the special emotional tendencies of Song poetry. "Guiqu" (returning home) and "guilai" (coming back) frequently appear

in Song ci, reflecting the mobility and changeability of Song dynasty society. The Song dynasty was a highly developed era of commodity economy, the peak of ancient Chinese commodity economy development, and a period of frequent social changes and mobility, where many people left their hometowns to seek a living, become officials, do business, or explore new territories, greatly enhancing social mobility. Among them, poets who could not tolerate political corruption and darkness yearned to retire to the countryside, or missed their hometowns and looked forward to the day of returning home, and wrote "guiqu" and "guilai" in their poems.

	唐朝关键词	词频	宋朝关键词	词频
11	惆怅	802	江南	696
12	使君	783	如今	622
13	秋风	760	回首	607
14	悠悠	742	阑干	603
15	相思	738	多少	590
16	如何	715	千里	589
17	青山	693	菩萨	581
18	白日	656	明月	558
19	何人	642	万里	528
20	相逢	641	芳草	506
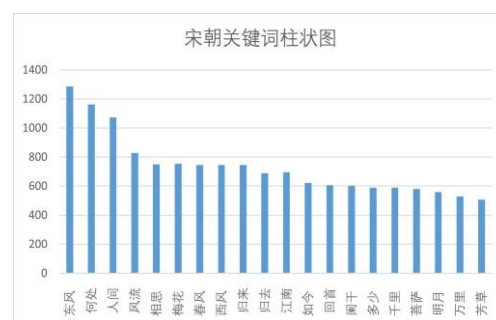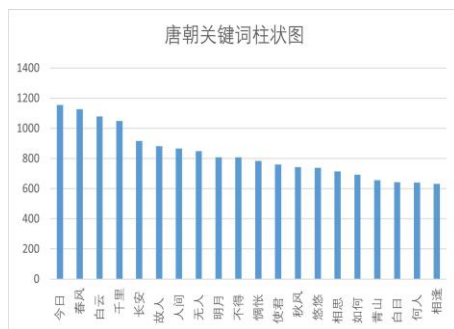
In Tang poetry, "autumn wind" is used to convey emotions of departure, homesickness, and the passage of time. It also gives a gentle force in solitude. For example, Li Bai's "Midnight Autumn Song" says "The autumn wind never ends, it always brings back memories of the Jade Pass." The persistent autumn wind carries deep feelings of missing those on the border, conveying the continuity of life. "Green mountains" are portrayed in Tang poetry as magnificent and grand, expressing philosophical metaphors of life. "White day" is generally depicted as the passing of time, the finite nature of life, and the fickleness of fortune, expressing the limited feeling of the author's lifetime. In Song lyrics, "the bright moon" often symbolizes the pursuit of beauty and yearning for good things, with romantic, pure, and lofty spiritual meanings. For example, Su Shi's "Water Tune Song Head" says "When will the bright moon appear? Let us raise our cups and ask the blue sky." This expresses the poet's emotional feelings of parting and heart's desires. Similarly, Xin Qiji's "Green Jade Case ・ Moon Festival" says "Where am I when I wake up tonight with wine? By the Yangliu River, enjoying the breeze and the lingering moonlight." Here, the bright moon represents the coexistence of beauty and sadness in the night. The word "fragrant grass", which appears frequently, is often used as a metaphor for good qualities or virtuous people. As such, it expresses aesthetic elegance and literary interest of literati culture. This shows that the Tang Dynasty valued the exploration of philosophy of life, focusing on everything present, natural and the self, adhering

to a free and unrestrained lifestyle and attitude. The Song Dynasty emphasized the pursuit of spiritual morality, which was closely related to the culture of scholar-officials. The pursuit of the "Yangchun Baixue" (literally means "the white snow of spring") elegant culture by the scholar-official class is fully reflected in Song poetry. The high frequency of appearance of scenery words in Song lyrics also demonstrates their elegance and literary interest.

"Chouchang" is the most common word in the 11-20 word phrases in Tang poetry, often used to express the transience of life and the fleeting nature of time. In Song lyrics, "Jiangnan" refers to the southern region of China, which is a significant geographical concept in Song poetry. Song lyricists express their memories of hometowns, homes, and peaceful living through borrowed scenery, revealing a sense of melancholy associated with specific periods. Based on the most commonly used words in Tang and Song poetry, "sorrow" is the prevailing theme throughout both eras, but it has different levels of depth. Throughout the Tang period, sorrow tended to be more superficially personal, stemming from feelings of sadness and frustration from an individual's life, expressing the author's loneliness, helplessness, and melancholy. However, during the Song dynasty, sorrow became more of a "social sorrow" and "national sorrow," stemming from changes in social environments. Sorrow was caused by concerns and anxieties about the political environment and social conditions, as well as the grief arising from societal upheaval, broken mountains and rivers, and wandering in distant lands. The resulting sense of national responsibility and mission is also reflected in poetic literature.

In terms of interrogative words and words indicating quantity, there are differences between Tang and Song poetry. Tang poetry uses fewer interrogative words, with the most frequently appearing being "how" and "who". On the other hand, Song poetry tends to use exaggerated words such as "thousands of miles", "ten thousand miles", and "how many", conveying a sense of vagueness and hopelessness. The emotional color of these words reflects the influence of the times and social environment on language style and emotional atmosphere. In summary, the differences between Tang and Song poetry reflect the development and evolution of ancient Chinese literature, as well as the influence of the times and social environment on culture and language.

Conclusion: Through our research in this paper, we have discovered significant differences in the high-frequency keywords of Tang and Song poetry. The high-frequency keywords in Tang poetry mainly cover natural landscapes, love, and homesickness, while the high-frequency keywords in Song poetry focus more on human relationships, social customs, and the overall societal landscape. This difference not only reflects the changes in social background and cultural atmosphere between the Tang and Song dynasties but also demonstrates the differences in artistic styles between the two eras. Through our study of Tang and Song poetry, we can gain a better understanding of the development and evolution of ancient Chinese literature, as well as a better understanding of the diversity and complexity of Chinese culture. Moreover, the Python program used in this research provides us with a new method for text analysis, offering new ideas and approaches for literary research and humanities and social science studies.