

# Project Gutenberg Analysis

Shinjini Ghosh

July 10, 2020

## Analysis

1. *Title of the novel analysed:* The Mysterious Affair at Styles by Agatha Christie
2. *Total number of words:* 56410
3. *Total number of unique words:* 9925
4. *Top 20 most frequent words:*

No.	Word	Count	No.	Word	Count
1	the	2624	11	in	821
2	i	1654	12	he	718
3	to	1336	13	is	538
4	of	1261	14	not	511
5	a	1201	15	had	501
6	and	1124	16	she	482
7	it	955	17	her	462
8	that	938	18	his	416
9	was	926	19	at	394
10	you	848	20	but	394

Table 1: Top 20 most frequent words

5. *Top 20 most frequent interesting words:* Here, we have excluded the top 1000 English words from consideration.

No.	Word	Count	No.	Word	Count
1	poirot	357	11	howard	83
2	inglethorp	244	12	asked	80
3	mrs	231	13	inglethorp's	72
4	mr	179	14	mary	69
5	john	161	15	dorcas	69
6	cavendish	137	16	looked	66
7	sir	106	17	strychnine	65
8	cynthia	98	18	alfred	64
9	into	95	19	dr	61
10	lawrence	84	20	evidence	56

Table 2: Top 20 most frequent interesting words