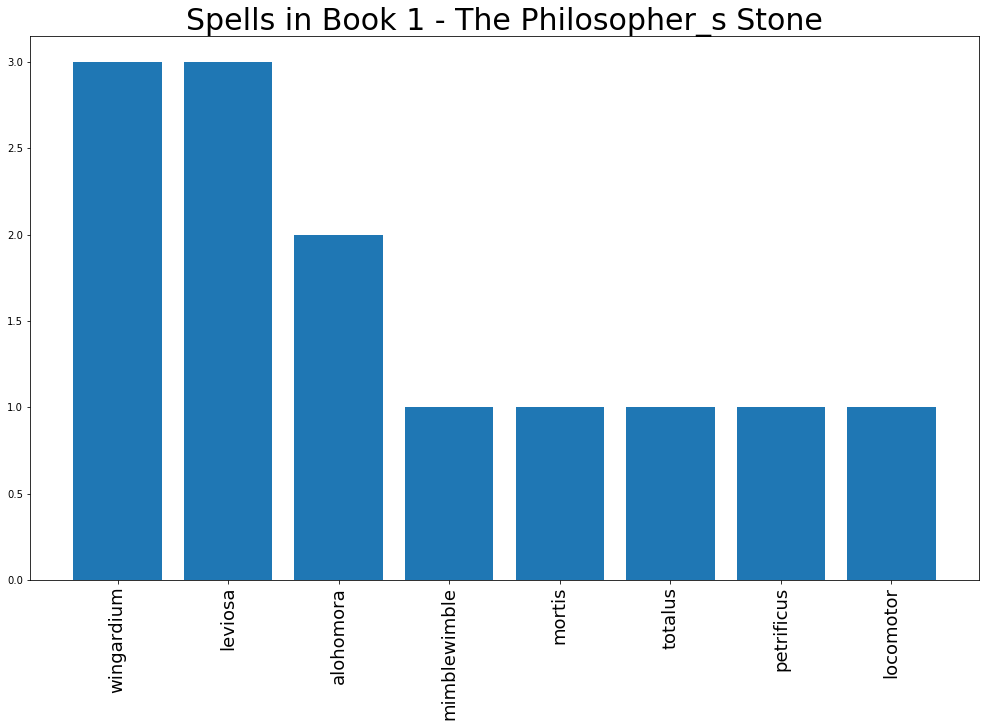
**Spell analysis**

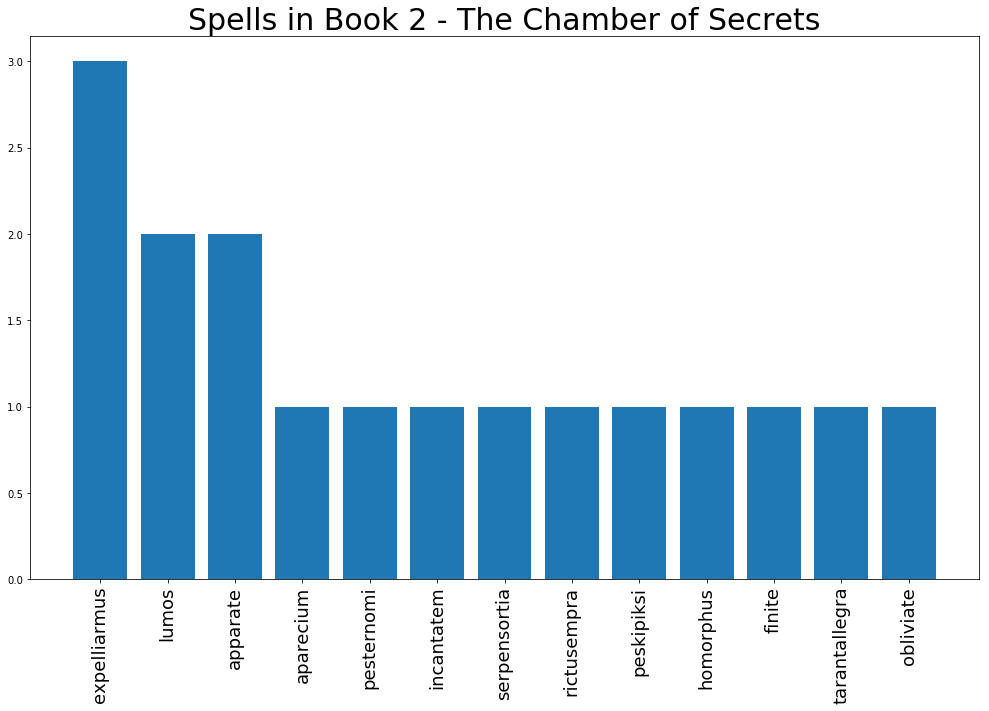
The frequency of each casted spell or curse was analyzed. In order to realize this we tokenized the books and then ran the frequency distribution function. Thereafter we matched the spell list excel sheet with each tokenized term count book. Tokenization is for lexical analysis, the process of classifying characters into individual words based on a space delimiter. This resulted in the frequency of spells and curses demonstrated below in a bar plot.

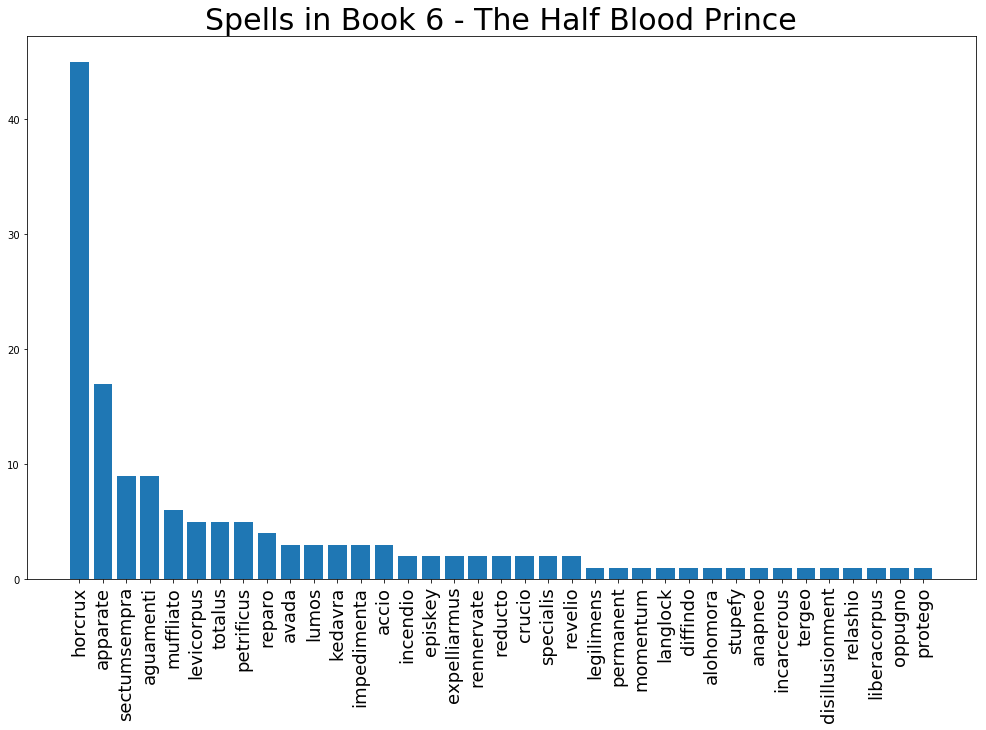
We expected given the nature of the story that the frequency would increase as the universe became more complex with each year. This turned out to be the case.

< Bar plots >



In the first book only few spells are casted and two curses, Mortis and Mimblewimble

In the second book we already have more spells but a good example is the 6th book where a lot of spells are casted and many curses.



As one can see then the horcrux curse is mentioned over 40 times. That is of no surprise since the plot of the book is the search for the seven horcruxes that Voldemort divided his soul into.

"Yes, I think so," said Dumbledore. "Without his

Horcruxes, Voldemort will be a mortal man with a

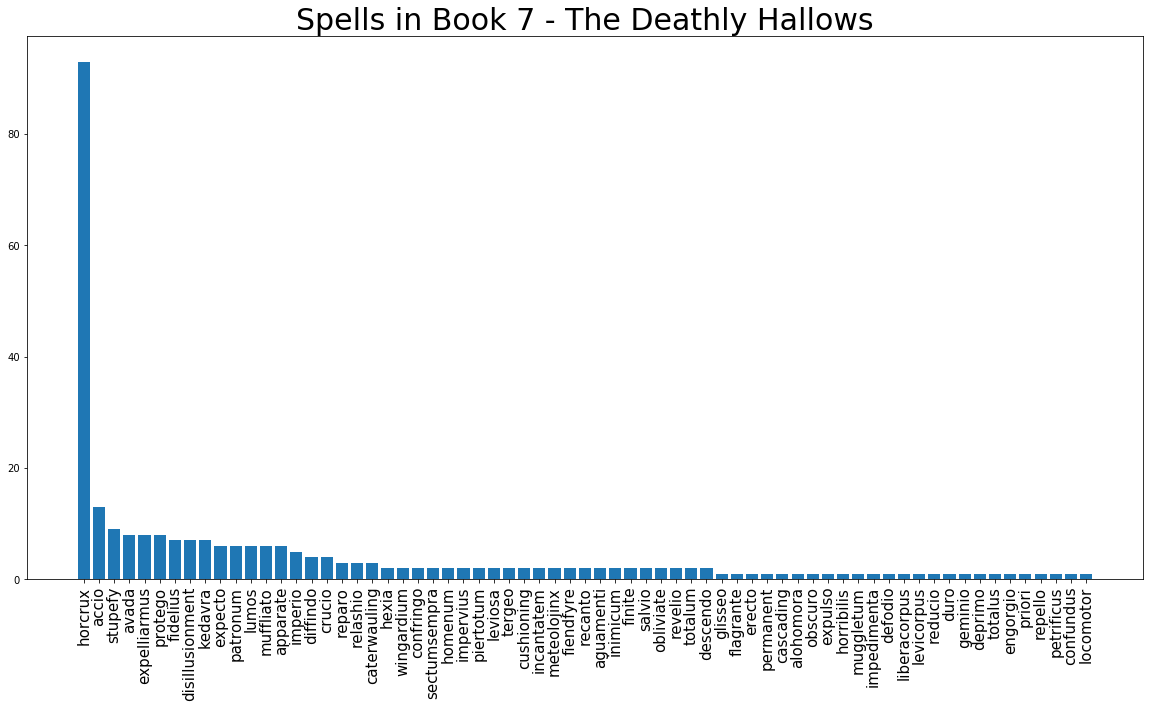
maimed and diminished soul. Never forget, though,

that while his soul may be damaged beyond repair,

his brain and his magical powers remain intact. It will

take uncommon skill and power to kill a wizard like

Voldemort even without his Horcruxes."

And finally we reach book 7 we can see that the search continues for the horcrux, appearing now over 80 times. Furthermore we now have the largest collection of spells and curses.

**TF-IDF**

We want to find the most representing words for each book and visualize them in a figure, a word cloud figure. In a word cloud figure the words differ is size based on how representing they are for a given book.

To realize this we utilize the TFIDF criterion

* TF: Term frequency, in each book see how often each word appears
* IDF: Inverse document frequency, see how important this word is compared to its appearance in other books. In other words to see if a given frequent word also appears in the other books.

The result is a TFIDF matrix, in our case the rows are corresponding to the books and the columns to the words.

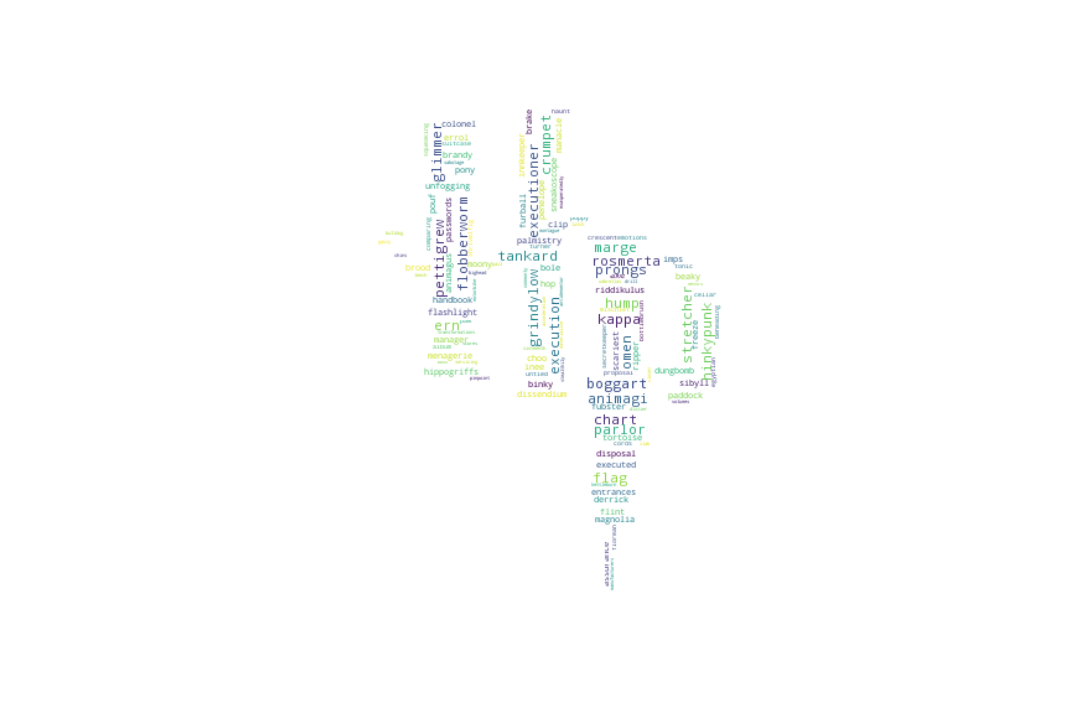
With this matrix we produce the seven word cloud figures, one for each book.

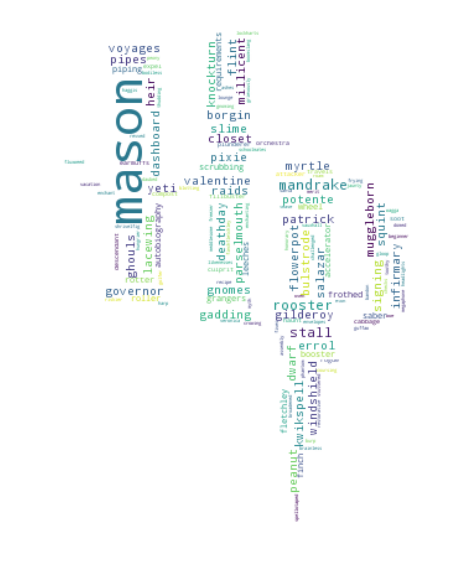
--------------------------------

**Figure for book 1**  
  


smelting : The school Dudley goes to  
bathrobe  
flamel: "Nicolas Flamel," she whispered dramatically, "is the only known maker of the Sorcerer's Stone!"  
mom :Mrs. Weasley, Ron’s mother is often referred to in this book  
pier

--------------------------------------------------------------------------------

Figure for book 2



mason: Mrs. Mason is they lady harry blows up in the air in the beginning of the book. Resulting in a dramatic discussion whether he should be kicked out of Hogwarts or not.



Figure for book 3

boggart :A boggart is an amortal shape-shifting non-being that takes on the form of the viewer's worst fear. The boggart was a big topic of this book since Harry's worst fear was a dementor at the time.  
Pettigrew: Is introduce in this book and the reader is informed that he in fact has been Rons rat  
grindylow : A Grindylow is a small, horned, pale-green water demon and a Dark creature  
hinkypunk :*The Hinkypunk is a little one-legged creature and looks as if it is made of smoke*  
executioner: An executioner from the Ministry of Magic that is sent to kill Buckbeak, Buckbeak turns out to be too valuable for Harry and the crew so Hermione helps the crew to travel back in time and save Buckbeak.

From the 5 five most representing words of this book, 3 of them are creatures from Professor Remus Lupin Defence against the dark art class, namely the boggart,grindylow and hinkypunk.

figure for book 4 

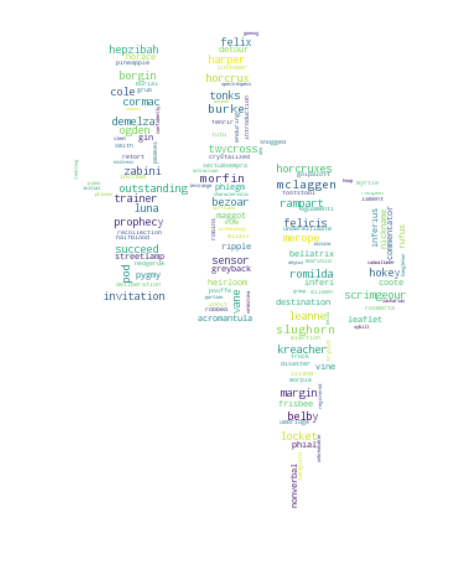
portkey: Is commonly used in this book. A portkey instantly brings anyone touching it to a specific location.  
Bulgarian: This book features both the world cup in Quidditch and the competition around the Goblet of fire where the Bulgarian team plays a big role.  
Karkaroff : Professor Karkaroff is the headmaster of Scandinavian magic school that the Bulgarian Viktor krum attends.  
Hostage: In the games of the goblet of fire then a hostage had to be saved  
leprechaun: Are magical creature that appear in this book

figure for book 5

bowtruckle Is a magical creature studied in the class care of magical creatures  
Thestrals : Skeleton flying horses, they were used in this book to fly the Orders member disguised as harry to a safe spot  
clipboard:  
umbridge: Dolores Jane Umbridge becomes a headmaster at hogwards in this book replacing dumbeldore.  
dai

-----------------------------------------------------------------------------------------------------

figure for book 6



Not really a word that represents this book

------------------------------------------------------------------------------------------------

figure for book 7



cattermole: Is the character that Ron is disguised as when the crew goes undercover into the ministry of magic  
xenophilius: He is the father of Luna Lovegood  
carrow Siblings that are both Death Eaters  
deluminator : Albus dumbledore's device which is used for taking out any light source and keeping them until bestowing them.   
Peverell: the Peverell family was the original owners of the Hallows. Antioch: the Elder Wand, Cadmus: the Resurrection Stone, and Ignotus: the Cloak of Invisibility