Analysis of Stephen King novels
May 2022

Prepared by Julia Kuznetsova

Scope of analysis

The analysis includes **63 novels** written by Stephen Kings and published as of May 2022.

1974	Carrie	1984	Thinner	1997	Wizard and Glass	2012	The Wind Through the Keyhole
1975	'Salem's Lot	1986	It	1998	Bag of Bones	2013	Joyland
1977	The Shining	1987	The Drawing of the Three	1999	The Girl Who Loved Tom Gordon	2013	Doctor Sleep
1977	Rage	1987	Misery	2001	Dreamcatcher	2014	Mr. Mercedes
1979	The Long Walk	1987	The Tommyknockers	2001	Black House	2014	Revival
1979	The Dead Zone	1989	The Dark Half	2002	From a Buick 8	2015	Finders Keepers
1980	Firestarter	1990	The Stand	2003	Wolves of the Calla	2016	End of Watch
1981	Roadwork	1991	The Waste Lands	2004	Song of Susannah	2017	Gwendy's Button Box
1981	Cujo	1991	Needful Things	2004	The Dark Tower		
1982	The Running Man	1992	Gerald's Game	2005	The Colorado Kid	2017	Sleeping Beauties
1982	The Gunslinger	1992	Dolores Claiborne	2006	Cell	2018	The Outsider
1983	Christine	1994	Insomnia	2006	Lisey's Story	2018	Elevation
1983	Pet Sematary	1995	Rose Madder	2007	Blaze	2019	The Institute
1983	Cycle of the Werewolf	1996	The Green Mile	2008	Duma Key	2021	Later
1984	The Talisman	1996	Desperation	2009	Under the Dome	2021	Billy Summers
1984	The Eyes of the Dragon	1996	The Regulators	2011	11/22/63	2022	Gwendy's Final Task

Method of comparison

63 novels were compared among themselves using **TfidfVectorizer,** a sklearn feature extraction tool that converts a collection of raw documents to a matrix of TF-IDF (term frequency — inverse document frequency) features.

Cosine similarity measure was used to compare features. For defining it, the sequences are viewed as vectors in an inner product space, and the cosine similarity is defined as the cosine of the angle between them. The cosine similarity always belongs to the interval [-1, 1], where **1 means the maximum similarity.**

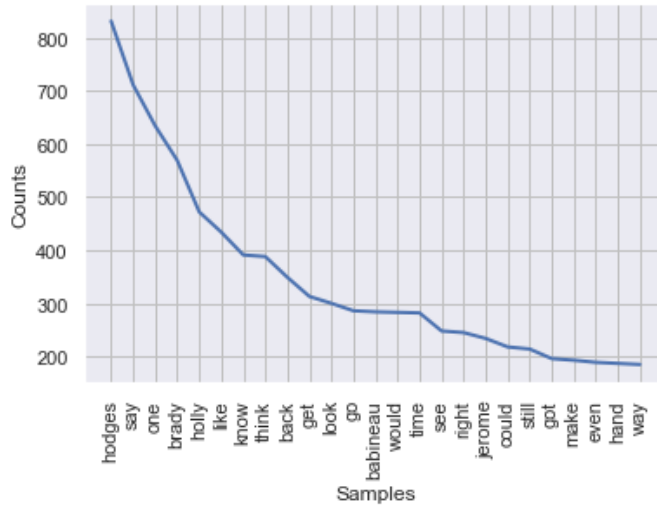**Similarity** means similarity in terms of words, word frequency, and phrasing.

The most similar novels

All novels in pairs of very similar books (similarity ≥ 0.95) **were written after 2009, except Bag of Bones.**
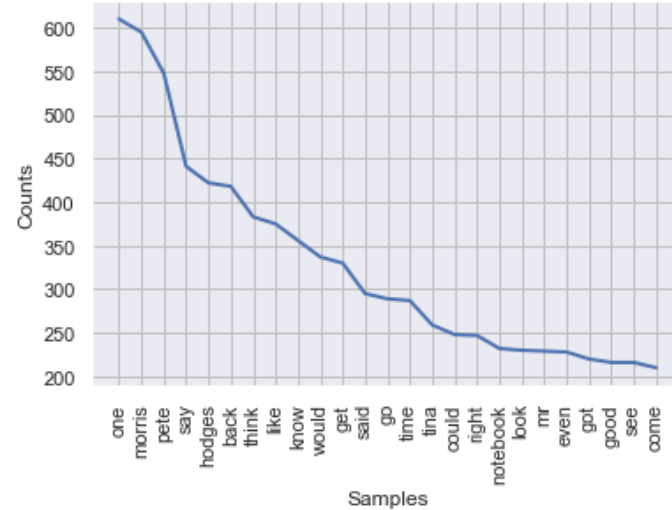
			Similarity
Pair 1	End Of Watch (2016)	Finders Keepers (2015)	0.966924
Pair 2	Gwendy's Button Box (2017)	Gwendy's Final Task (2022)	0.965771
Pair 3	Joyland (2013)	Under The Dome (2009)	0.964749
Pair 4	Joyland (2013)	Revival (2014)	0.956291
Pair 5	Joyland (2013)	Bag Of Bones (1998)	0.951275
Pair 6	Joyland (2013)	The Outsider (2018)	0.950588
Pair 7	Revival (2014)	Under The Dome (2009)	0.950219

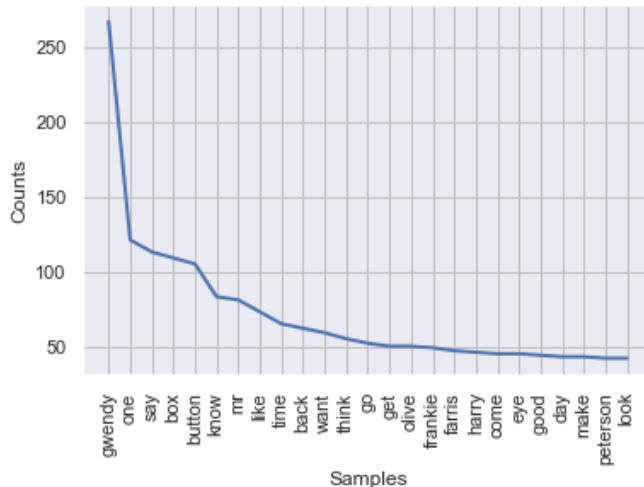# The most similar novels: word frequency

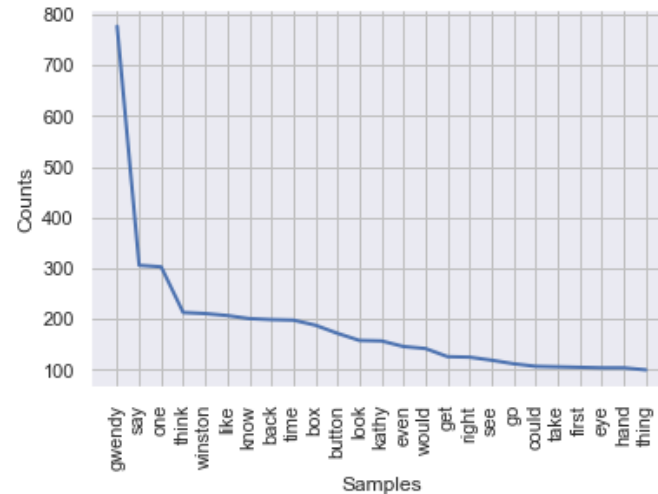Word frequency in the novel End Of Watch

Word frequency in the novel Finders Keepers

Word frequency in the novel Gwendys Button Box

Word frequency in the novel Gwendys Final Task

The least similar novels

The Gunslinger and The Long Walk are very different (similarity ≤ 0.05) from several novels.

			Similarity
Pair 1	The Gunslinger (1982)	End Of Watch (2016)	0.040495
Pair 2	The Gunslinger (1982)	Finders Keepers (2015)	0.049840
Pair 3	The Long Walk (1979)	End Of Watch (2016)	0.043786
Pair 4	The Long Walk (1979)	Gwendy's Button Box (2017)	0.042023

Similarity among The Dark Tower novels

The most similar among The Dark Tower novels are:

- The Wind Through The Keyhole and The Dark Tower (similarity = 0.916445)
- The Waste Lands and Wolves Of The Calla (similarity = 0.868579
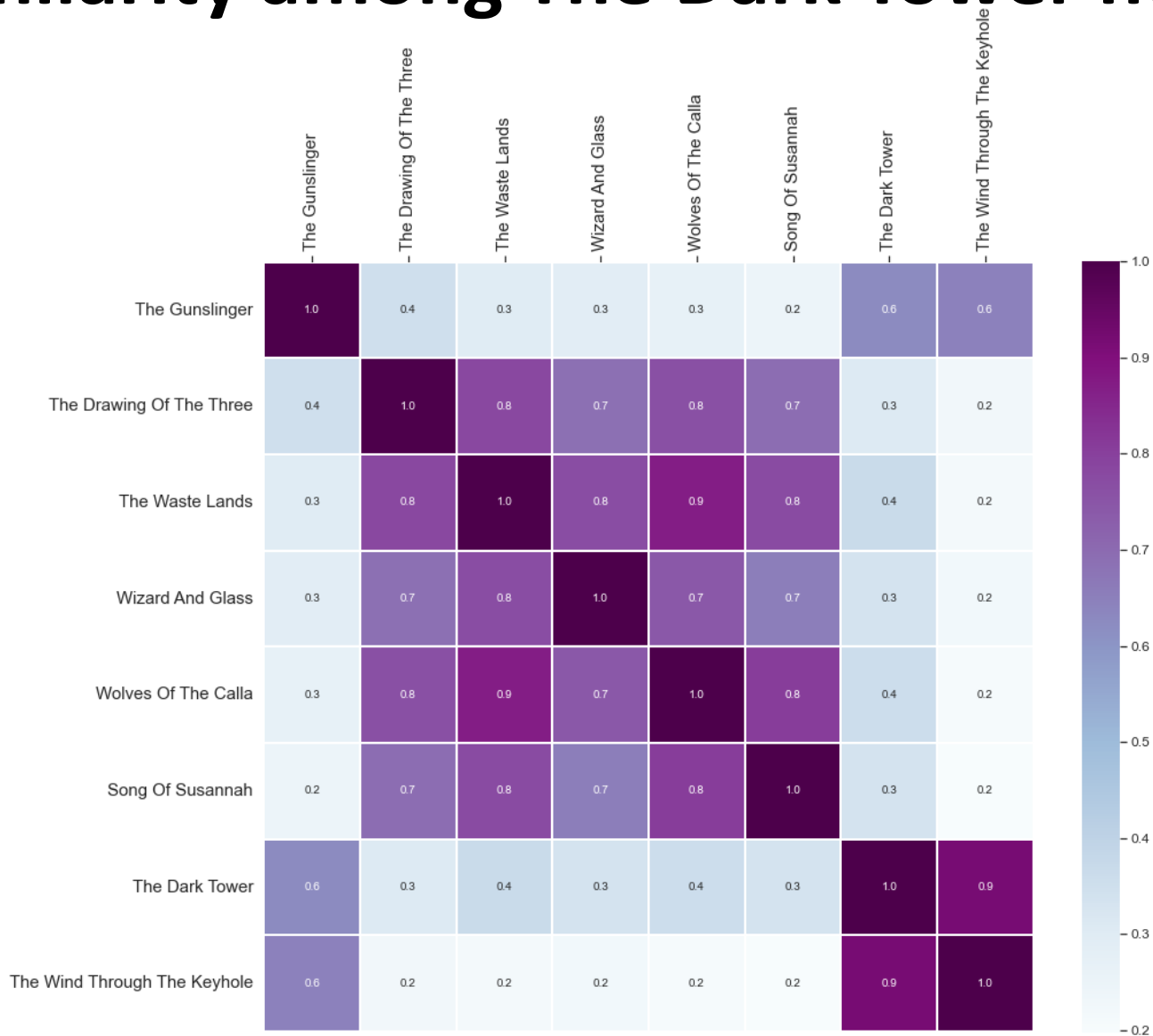- Song Of Susannah and Wolves Of The Calla (similarity = 0.807468)

The Wind Through The Keyhole is **the most different** among novels of this series:

- The Wind Through The Keyhole and Song Of Susannah (similarity = 0.197719)
- The Wind Through The Keyhole and Wolves Of The Calla (similarity = 0.209151)
- The Wind Through The Keyhole and The Waste Lands (similarity = 0.217377)
- The Wind Through The Keyhole and Wizard And Glass (similarity = 0.225490)
- The Wind Through The Keyhole and The Drawing Of The Three (similarity = 0.226433)

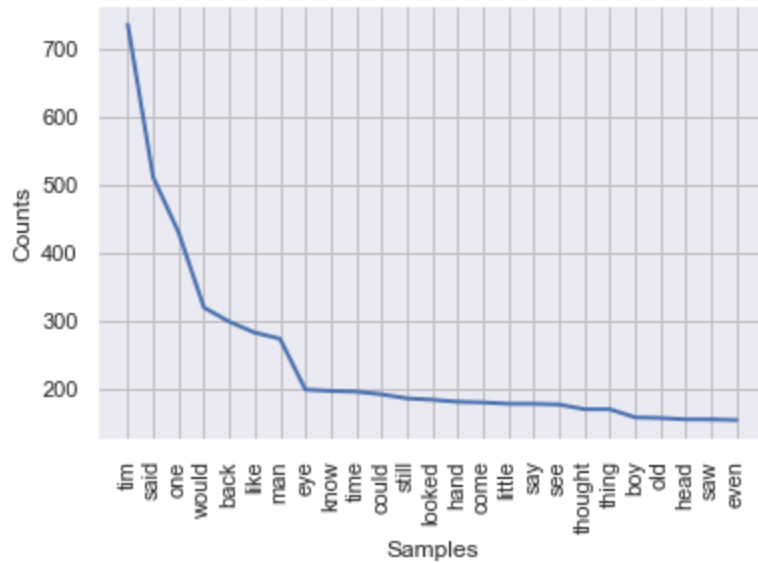Thus, there are **two groups of The DT novels united by word similarity:**

- The Gunslinger, The Dark Tower, and The Wind Through The Keyhole
- The rest (The Drawing Of The Three, The Waste Lands, Wizard and Glass, Wolves Of The Calla, and Song of Susannah)

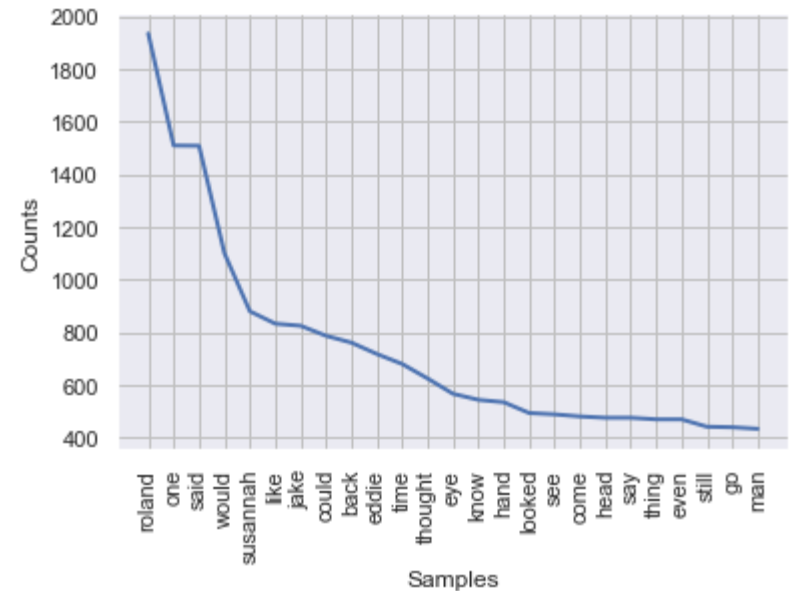# Similarity among The Dark Tower novels

Word frequency in The Dark Tower novels

Word frequency in the novel The Wind Through The Keyhole

Word frequency in the novel The Dark Tower

Novels written in co-authorship

Five novels were written in co-authorship:

- The Talisman (1984, with Peter Straub)
- Black House (2001, with Peter Straub)
- Gwendy's Button Box (2017, with Richard Chizmar)
- Sleeping Beauties (2017, with Owen King)
- Gwendy's Final Task (2022, with Richard Chizmar)

Similarity analysis has shown that:

- The Talisman and Black House are different from other Stephen King books (similarity < 0.5)
- Both Gwendy novels are very similar between each other (similarity = 0.96)
- Sleeping Beauties wording looks very similar (≥ 0.85) to a variety of Stephen King books
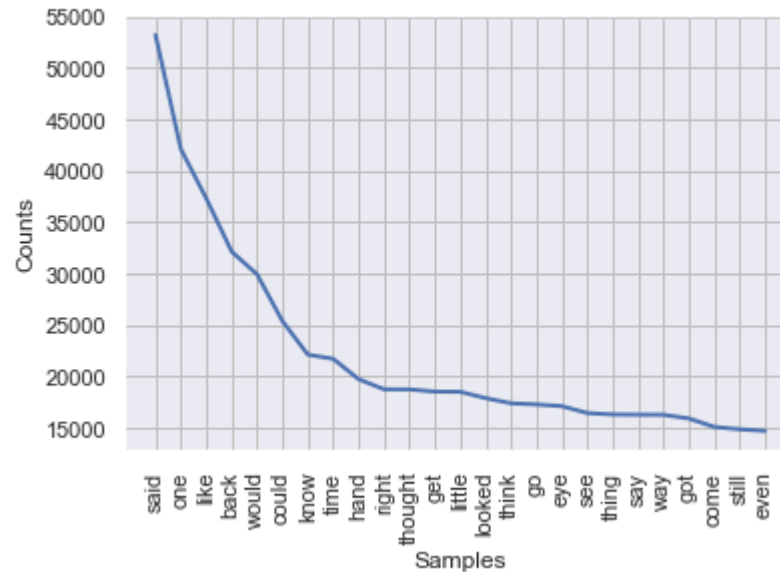
Novels similar to Sleeping Beauties

The following novels are similar to Sleeping Beauties (similarity ≥ 0.85):

Joyland (2013)	0.949744	The Wind Through The Keyhole (2012)	0.914879
Under The Dome (2009)	0.945271	The Colorado Kid (2005)	0.913519
Revival (2014)	0.938211	Cell (2006)	0.913498
The Outsider (2018)	0.935828	Billy Summers (2021)	0.911454
Bag Of Bones (1998)	0.933672	The Regulators (1996)	0.910473
Needful Things (1991)	0.928754	The Dark Tower (2004)	0.910413
Later (2021)	0.928649	Elevation (2018)	0.907241
Duma Key (2008)	0.928171	Blaze (2007)	0.895995
Doctor Sleep (2013)	0.923030	Mr Mercedes (2014)	0.887420
The Institute (2019)	0.916117	Insomnia (1994)	0.874596

Word frequency in all Stephen King novels

Word frequency in all Stephen King novels

Contacts

- **GITHUB:** github.com/kooznitsa
- **EMAIL:** jukuznetswork@gmail.com