Birth of the Modern Detective Story: Data dictionary

## General notes

* Variables with all missing have been removed from the data. That is, you may see options in the data collection information that do not appear in this dataset because no stories had any information related to that option.
* There are four variables that are about the number of words:
  + text\_words count how many words in the story altogether
  + before\_reveal\_words counts how many words before the ‘reveal border sentence’
  + after\_reveal\_words counts how many words after the reveal border sentence
  + title\_words counts how many words in the title/name of the story

Other than title\_words, these these variables have quite a few missing values, due to not having access to the text for the stories in a scrapable format. They are also potentially somewhat unreliable as disclaimer or other information text at the beginning/end may have been captured, e.g., “This work is out of copyright in countries with a copyright period of 70 years or less, after the year of the author's death.” We may be able to provide more (and more accurate results) in the when your receive the data, but don’t *count* on it.

| Variable name | Data type | Levels / summary | Number of obs missing |
| --- | --- | --- | --- |
| story\_code | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 0 |
| story\_name | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 0 |
| plot\_summary\_25\_50\_words | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 0 |
| author\_code | character | "ACD", "AM", "BO", "MPS", "CKP", "JF", "EB", "AF", "EWH", "JEPM", "ABR", "AKG", "VW", "IZ", "RB", "MDP", "TWH", "WHH" | 0 |
| in\_what\_order\_are\_the\_investigation\_and\_reveal\_portions\_provided | character | "Standard", "Reverse", "Other" | 0 |
| name\_of\_detective\_number\_1 | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 0 |
| detective\_number\_1\_gender | character | "Male", "Female", "Unknown" | 0 |
| detective\_number\_1\_role | character | "Independent amateur", "Independent professional (solo)", "Independent professional (agency)", "Member of the police", "Unclear" | 0 |
| name\_of\_detective\_number\_2 | character | "Louis Carlyle", "Hutchinson Hatch", "\_unknown" | 350 |
| detective\_number\_2\_gender | character | "Male" | 350 |
| detective\_number\_2\_role | character | "Independent professional (agency)", "Independent amateur", "Independent professional (solo)" | 350 |
| name\_of\_assistant\_number\_1 | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 100 |
| assistant\_number\_1\_gender | character | "Male", "Female", "Unknown" | 100 |
| assistant\_number\_1\_role | character | "Independent amateur", "Independent professional (solo)", "Independent professional (agency)", "Unclear" | 100 |
| name\_of\_assistant\_number\_2 | character | "Shinwell Johnson", "Billy", "Parkinson", "Christopher Jervis", "Podger", "Narkom" | 343 |
| assistant\_number\_2\_gender | character | "Male" | 343 |
| assistant\_number\_2\_role | character | "Independent amateur", "Independent professional (agency)" | 343 |
| are\_one\_or\_more\_culprits\_introduced\_into\_the\_story\_only\_during\_the\_reveal | character | "No", "Yes" | 13 |
| the\_main\_culprit | character | "Acts mostly alone with some assistance", "Acts alone", "Acts together with one or more equal partners", "Acts on behalf of a crime syndicate" | 15 |
| who\_initiates\_the\_detectives\_investigation | character | "The police", "A private client", "The assistant(s)", "The detective(s) themselves", "Unclear/not specified", "The detective(s) themselves, Unclear/not specified", "The police, A private client", "A private client, The detective(s) themselves", "A private client, The assistant(s)" | 0 |
| what\_is\_the\_role\_of\_the\_police\_force\_in\_solving\_the\_crime | character | "Police are present but with no role, or only a minor role, in helping solve the crime", "Police are not present", "Police are present and contribute significantly to solving the crime", "Police are present and actively interfere with or hinder the solution of the crime" | 1 |
| the\_storys\_main\_action\_focuses\_on\_a | character | "Crime", "Quasi-crime" | 10 |
| when\_do\_es\_the\_crime\_s\_mystery\_being\_investigated\_occur | character | "Before the investigation", "Before the investigation, During the investigation", "Before the investigation, After the investigation", "During the investigation", "Before the investigation, During the investigation, After the investigation", "During the investigation, After the investigation", "After the investigation" | 1 |
| types\_of\_crimes\_or\_quasi\_crimes\_present\_in\_story | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 10 |
| motives | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 11 |
| means\_murder\_only | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 152 |
| does\_the\_story\_provide\_sufficient\_clues\_in\_sufficient\_detail\_to\_allow\_an\_alert\_reader\_to\_correctly\_guess\_the\_solution\_before\_the\_reveal | character | "Yes", "No" | 11 |
| does\_the\_story\_provide\_sufficient\_clues\_in\_sufficient\_detail\_to\_allow\_an\_alert\_reader\_to\_definitively\_solve\_the\_crime\_before\_the\_reveal | character | "No", "Yes" | 11 |
| did\_either\_annotator\_personally\_correctly\_guess\_the\_solution\_to\_the\_crime\_before\_the\_reveal | character | "No", "Yes" | 11 |
| types\_of\_clues | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 5 |
| what\_is\_the\_essential\_clue\_in\_the\_story\_if\_present | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 19 |
| of\_what\_type\_is\_the\_essential\_clue\_if\_present | character | "Physical clue", "Testimonial clue" | 19 |
| what\_is\_the\_most\_salient\_clue\_in\_the\_story\_if\_clues\_are\_present | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 19 |
| of\_what\_type\_is\_the\_most\_salient\_clue\_if\_clues\_are\_present | character | "Physical clue", "Testimonial clue" | 19 |
| describe\_the\_main\_red\_herring\_if\_present | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 285 |
| of\_what\_type\_is\_the\_main\_red\_herring\_if\_present | character | "Testimonial clue", "Physical clue" | 285 |
| presence\_of\_planted\_or\_fabricated\_evidence | character | "Fabricated", "Planted", "Planted, Fabricated" | 243 |
| types\_of\_evidence\_made\_available\_in\_the\_story | character | "Testimonial evidence, Physical evidence", "Testimonial evidence", "Physical evidence" | 6 |
| is\_the\_crime\_solved | character | "Yes", "No" | 0 |
| does\_the\_detective\_pointedly\_and\_independently\_decide\_not\_to\_alert\_the\_authorities\_to\_the\_facts\_of\_the\_crime\_despite\_the\_fact\_that\_an\_actionable\_crime\_has\_been\_committed | character | "Yes", "No" | 2 |
| do\_you\_receive\_a\_satisfying\_narrative\_of\_account\_of\_the\_crime\_s\_in\_all\_their\_relevant\_details\_and\_a\_summary\_of\_the\_process\_by\_which\_the\_crime\_s\_are\_solved | character | "Yes", "No" | 0 |
| would\_you\_recommend\_this\_story\_to\_a\_friend | character | "Yes", "No" | 1 |
| crime\_trajectory | character | "Qrime → different qrime", "No qrime → no qrime", "No qrime → qrime", "Qrime → same qrime", "Qrime → no qrime" | 2 |
| what\_is\_the\_reveal\_border\_sentence | character | Too many long/unique levels to show here. See the data collection instructions for more information. | 9 |
| surname\_s | character | This is about the author. "Concan Doyle", "Morrison", "Orczy", "Shiell", "Pirkis", "Futrelle", "Bramah", "Freeman", "Hornung", "Muddock", "Reeve", "Green", "Whitechurch", "Zangwill", "Barr", "Post", "Hanshew", "Hodgson" | 0 |
| given\_name\_s | character | This is about the author. "Arthur", "Emma", "Matthew Phipps", "Catherine Louisa", "Jacques", "Ernest", "Austin", "Ernest William", "J. E. Preston", "Arthur B.", "Anna Katherine", "Victor", "Israel", "Robert", "Melville Davisson", "Thomas W.", "William Hope" | 0 |
| sex | character | This is about the author. "M", "F" | 0 |
| country\_of\_birth | character | This is about the author. "Scotland", "England", "Hungary", "Montserrat", "United States of America" | 0 |
| city\_town\_of\_birth | character | This is about the author. "Edinburgh", "London", "Tarnaörs", "Pike County, Georgia", "Manchester", "Cleveland Villas", "Southampton", "Brooklyn, New York", "Norham, Northumberland", "Glasgow", "Harrison County, West Virginia", "Blackmore End, Essex" | 3 |
| country\_of\_death | character | This is about the author. "England", "Henley-on-Thames", "Chichester", "Atlantic Ocean", "France", "United States of America", "Belgium" | 0 |
| city\_town\_of\_death | character | This is about the author. "Crowborough", "Chalfont St. Peter", "England", "London", "RMS Titanic", "Gravesend", "Saint-Jean-de-Luz", "Trenton, New Jersey", "Buffalo, New York", "Buxton, Debyshire", "Midhurst, West Sussex", "Woldingham, Surrey", "Harrison County, West Virginia", "Ypres salient" | 0 |
| nationality\_ies | character | This is about the author. "British, Scottish", "British, English", "Hungarian, British", "Caribbean, British", "American" | 18 |
| viaf | character | This is about the author. "https://viaf.org/viaf/65283845/", "https://viaf.org/viaf/27069152/", "https://viaf.org/viaf/39382131/", "https://viaf.org/viaf/54194618/", "https://viaf.org/viaf/77779811/", "https://viaf.org/viaf/9880305/", "https://viaf.org/viaf/76564127/", "https://viaf.org/viaf/125144814461944904157/", "https://viaf.org/viaf/71409962/", "https://viaf.org/viaf/26949841/", "https://viaf.org/viaf/73002380/", "https://viaf.org/viaf/69054865/", "https://viaf.org/viaf/76448216/", "https://viaf.org/viaf/56617927/", "https://viaf.org/viaf/122140022/", "https://viaf.org/viaf/129012558/", "https://viaf.org/viaf/24320742/", "https://viaf.org/viaf/24602763/" | 0 |
| story\_url | character | This is about the story.  Too many long/unique levels to show here. See the data collection instructions for more information. | 21 |
| format\_of\_first\_publication\_journal\_or\_book | character | This is about the story.  "Journal", "Book" | 26 |
| name\_of\_journal\_or\_title\_of\_book | character | This is about the story.  Too many long/unique levels to show here. See the data collection instructions for more information. | 26 |
| year\_of\_birth | Date | This is about the author.  Date variable | 0 |
| date\_of\_death\_yyyy\_mm\_dd | Date | This is about the author.  Date variable | 0 |
| date\_of\_first\_publication\_yyyy\_mm\_dd | Date | This is about the story.  Date variable | 21 |
| more\_than\_two\_assistants\_if\_yes\_enter\_total\_number | numeric | Min: 3, Mean: 3, Median: 3, Max: 3 | 352 |
| number\_of\_victims\_of\_gender\_male | numeric | Min: 0, Mean: 1.6, Median: 1, Max: 100 | 0 |
| number\_of\_victims\_of\_gender\_female | numeric | Min: 0, Mean: 0.6, Median: 0, Max: 5 | 0 |
| number\_of\_culprits\_of\_gender\_male | numeric | Min: 0, Mean: 7.4, Median: 1, Max: 2000 | 0 |
| number\_of\_culprits\_of\_gender\_female | numeric | Min: 0, Mean: 0.5, Median: 0, Max: 3 | 0 |
| number\_of\_culprits\_of\_gender\_non\_binary | numeric | Min: 0, Mean: 0, Median: 0, Max: 1 | 0 |
| number\_of\_culprits\_of\_gender\_unknown | numeric | Min: 0, Mean: 0, Median: 0, Max: 3 | 0 |
| how\_satisfying\_is\_this\_story\_as\_a\_piece\_of\_detective\_fiction | numeric | Min: 1, Mean: 2.8, Median: 3, Max: 5 | 0 |
| text\_words | numeric | Min: 319, Mean: 36451.9, Median: 35040.5, Max: 73494 | 235 |
| before\_reveal\_words | numeric | Min: 1622, Mean: 28876.8, Median: 27087, Max: 57968 | 262 |
| after\_reveal\_words | numeric | Min: 1439, Mean: 9779.6, Median: 8134, Max: 35926 | 262 |
| title\_words | numeric | Min: 8, Mean: 24.1, Median: 22, Max: 67 | 0 |
| year\_published | numeric | Min: 1811, Mean: 1905.7, Median: 1905, Max: 1927 | 21 |