

Annotation Guidelines for Condensation Project

General Instructions:

Each datapoint consists of a New York Times article headline accompanied with its 2-3 sentence description. The datapoint must be ranked on a scale of 1 (Not condensing at all) to 5 (Very blatantly condensing) towards a given target.

Condensation Scale:

As briefly mentioned above, we have defined condensation on a scale between 1 to 5, where 1 is 'not condensing at all' and 5 is 'very blatantly condensing'.

Ranking	Description
1	The data entry is completely factual and is not trying to evoke any particular emotion but is simply stating what has happened, or is referring to a statement made by someone, regardless of whether that statement itself is condensing or not.
2	The data entry is adding a subtle commentary on the topic, likely with the use of a special word (ex. using "battered" instead of "bombed"). Indicates slight superiority over or sympathy towards the subject with these word/tone choices.
3	The data entry includes more than just one careful selection of a word with a condensing connotation; there are multiple references to that point of condensation and there is an ongoing theme of that condensation throughout the text. It is similar to a ranking of 2 but with repeated instances instead of a singular one.
4	The data entry illustrates clear commentary about the subject, almost like an opinion, but it is not intentionally trying to be rude to the subject. The key purpose in this case is to report on the event, not to be overtly condensing.
5	The data entry intentionally and blatantly refers to the subject in a condensing manner using punctuation, word choice, and opinionated language. The key purpose of the headline and description is to express condensation towards the subject.

Identifying the Target of Condescension

The purpose of this project is to test the NYT's condescension towards countries and individuals, thus condescension should be rated with regards to the foreign country/person/organization mentioned. Some guidelines:

- If the US and another country are mentioned, the annotator must judge the condescension towards the foreign country
- If no explicit foreign country is mentioned, but a representative of the country (E.g. the President) or an organization is mentioned, the annotator must judge condescension towards this target
- If multiple foreign countries are mentioned, the annotator must judge condescension towards the country that they judge to be the most relevant / the country that is the focus of the article

Hyperbolic Terms

We define hyperbolic terms to be ones that suggest exaggeration beyond that which is supported by the facts. These may often occur in the form of adverbs such as *Always*, *Never*, *Totally*, etc. The use of a hyperbolic term should cause the annotator to increase their condescension rating for the given data point as the writer is making assumptions about the target which suggests some superiority over them. Some examples can be seen below:

Shanghai introduces a staggered lockdown by district in an effort to test the entire city.

Unlike other Chinese cities, Shanghai had never before imposed a full-scale citywide lockdown, but the current measure comes close.

- The above example does not count as usage of hyperbolic terms because despite the usage of 'never', the sentence is still stating a fact.
- The example below counts as a usage of a hyperbolic term because the usage of 'never' is hyperbolic and causes the reader to feel pity for the villagers and believe they are unable to move on.
 - *"The village will never forget the thirst that plagues them"*

Author Making Claims

If the datapoint makes claims about the motivations or intentions of its target that is not backed up by data, then this should increase the condescension rank. In the following example, the claim that France doesn't care about its Green Party is the opinion of the author because "not caring" has not been backed up by data and is therefore, condescending toward France.

France Cares About Green Causes, but Not Its Green Party

As a presidential election looms, the Greens lag far behind in the polls. Analysts say the party has failed to inspire voters and show them it can rule.

As a side note, it is not necessary for the *entire* datapoint to be condescending to warrant a higher rank. In the above example, only the headline has the author's opinion whereas the description is factual and objective but since condescension occurs at all, this should cause the annotator to award a higher rank.

Identifying Tone

Tone can be discovered by carefully evaluating the choice the author makes about the specific words that are present in that text. Tone implies an idea of the 'other'; implying the situation described in the headline and title could never happen in other countries/territories or it indicates surprise of a country's/territory's accomplishments. This includes opinions that are not directly factual and show some bias on the part of the writer. Tone that makes the subjects of the articles feel like deserving of pity should also be considered. For instance, if the author is insinuating that they need saviors ("Hopeless", "Vulnerable"). This could also include describing a situation or attribute as a weakness when it may not be perceived as a weakness by the individual of focus (ex. Implying deafness is a weakness when it may not be viewed this way by deaf people).

Example:

'Improbable Journey': How a Movie From Tiny Bhutan Got an Oscar Nod

"Lunana: A Yak in the Classroom" was filmed on a shoestring budget in a remote Himalayan village. It's now an Academy Award nominee, a first for Bhutan.

The use of adjectives such as "tiny" to describe the country Bhutan and "shoestring" for the budget demeans the country and its accomplishment, making it seem that the Oscar nod was a farfetched dream for the country and those involved in the movie.

Citing Studies or Quotes

If a datapoint includes a quote from an individual, article or study, then even if that quote itself is condescending, it should **not** contribute to an increased condescension rank for the given datapoint. This is because the datapoint is simply reporting what has been said by someone/something else and in doing so, it is equivalent to a rating of 1 (since it's still factual). For example:

Ravaged by Cholera, Yemen Faces 2nd Preventable Scourge: Diphtheria
Aid officials said the Saudi blockade of Yemen was now impeding their ability to fight diphtheria

Since this was said by the Aid officials, it doesn't matter if what was said was condescending. However, keep in mind that there may still be condescension cues outside the quote elsewhere in the datapoint.

Evoking Emotion in the Annotator vs Unintentional Condescension

Just because the annotator may feel pity/anger/repulsion for a given subject after reading the datapoint does not necessarily mean that the datapoint itself was condescending. This could be because of their preexisting feelings about the target of the datapoint or their personal opinion on the topic being discussed in the article. To give a datapoint a condescension rank greater than 1, there must be explicit use of hyperbolism, condescending tone, or any of the other indicators mentioned in these guidelines. Example”

India Election Victory Is Expanding a Hindu Monk's National Profile
Yogi Adityanath's return as chief minister of Uttar Pradesh is fueling talk that he might succeed Narendra Modi as prime minister one day, and continue to advance their Hindu political movement

The annotator could have their own beliefs about whether the Hindu Nationalism politics prominent in India is good/bad for the country and this could make them feel this article is either defending/attacking it, which could come across as condescending. However, since the article is repeating facts, it cannot be marked as condescending for this.

The writer of the article may also be unintentionally condescending but despite the intentionality, this should still be used to increase the condescension rank of the datapoint.

Sri Lanka's Plunge Into Organic Farming Brings Disaster
The economically troubled country banned chemical fertilizers without preparing farmers, prompting a surge in food prices and worries about shortages.

The usage of “economically troubled” and “disaster” can be viewed as mildly condescending here and would be enough to warrant a rank of 2 because despite being unintentional, it is being hyperbolic and condescending towards Sri Lanka.

Pitfalls

- Uneven Distribution: it is perfectly valid to get an unequal distribution of ranks, most of the documents will likely be ranked with a 1, and there should generally be very few 4s and 5s. 2s should be relatively more common than 3s because it is likely that the condescension is more subtle in the form of word choice or a slight emotion.
- Convincing Journalism: Make sure to not be fooled by elements of convincing journalism, where authors are writing to capture the attention of the reader. This is a headline and description so it needs to serve the purpose of attracting readers, so make sure to not confuse that with condescension (especially in terms of rank 2 articles that have to do with word choices).
- Distinguishing between a 2 and a 3: A rank 2 has to do with a single word choice, and a very subtle implication. 3 has to do when that is repeated more than once and there is a tone being evoked here where the opinion of the author starts coming through but not in a deliberate way that is extreme because that would enter the rank 4 and 5 territory.
 - For instance, in the document below, we can see that there is a tone here of criticism being evoked and the use of the phrase “actual power” is more than just a single word being specifically chosen for a condescending purpose and there is a general theme making it more of a 3 than a 2.

Hong Kong Independence Movement Contests Beijing at the Polls

A new, youthful force, supporting some degree of greater self-determination for Hong Kong, may acquire a small measure of actual power in the election on Sunday.

Some Example Data Points with Advice:

- 1) Need to differentiate between empathy and sympathy for condescension. In the following example, the “*Trapped by War*” may be mistakenly viewed as condescension since it implies the helplessness of the victims and could be thought to generate sympathy. However, since this is the reality of the situation and the article isn’t being hyperbolic, it should receive a ranking of 1.

Link	Document	Ranking
https://www.nytimes.com/2022/03/12/world/europe/ukraine-surrogate-mothers-babies.html?searchResultPosition=21	<p>In a Kyiv Basement, 19 Surrogate Babies Are Trapped by War but Kept Alive by Nannies</p> <p>Some surrogate mothers are trapped by the fighting as their due dates near. And newborns face uncertain fates, with many biological parents now unable to travel to Ukraine.</p>	1

- 2) Even if the annotator feels the subject is deserving of condescension, they should still give a higher rank if it is present in the data point and not give it a high rank if it is merely stating a fact or reporting what happened. In the following article, there is a tone of condescension towards the Russian forces and despite what the annotator feels about the Russia-Ukraine conflict, this must still be noted. Similar cases may happen with articles noting terrorist groups, cruel politicians, etc.

Link	Document	Ranking
https://www.nytimes.com/2022/03/11/world/europe/ukraine-kyiv-russia-fighting.html?searchResultPosition=36	<p>How a Line of Russian Tanks Became an Inviting Target for Ukrainians</p> <p>A skirmish in an eastern suburb of Kyiv illustrated how close Russian forces have come as they continued to tighten a noose on the city,</p>	3

	but also revealed their strategic vulnerabilities.	
--	--	--

- 3) The connotations of the words they use, “bitter” recollections here makes this more than just a fact but is not enough to be a 3 because there isn’t a repetition of that or a specific tone carried out throughout the text.

Link	Document	Ranking
https://www.nytimes.com/2022/03/11/world/europe/syria-ukraine-war.html?searchResultPosition=39	Russian Sieges of Ukrainian Cities Provoke Bitter Recollections for Syrians “We feel for them and say, may God be with you,” said one Syrian who endured hunger and cold when his own government, backed by Russia, besieged Aleppo.	2

- 4) If there are extra adjectives that don’t really have anything to do with fact and insinuate an opinion about the subject, it is condescending. In the below example, “In Russia’s Shadow” is not necessary and doesn’t really do anything here other than demean the Baltics:

Link	Document	Ranking
https://www.nytimes.com/2022/03/07/world/europe/baltics-russia-ukraine-war-blinken.html?searchResultPosition=123	Baltics, in Russia's Shadow, Demand Tougher Stance From West “We have no illusions about Putin’s Russia anymore,” said one Baltic official in a meeting with the U.S. secretary of state.	2

- 5) Use of hyperbolic terms and sentence phrasing to focus on shortcomings or failings of the subject: The headline and description introduce the actresses’ so-called shortcomings first, focusing on the subject’s hardships as opposed to celebrating her accomplishments.

As a result, this condescending tone seems to demean the subject's past as opposed to celebrating or reporting on it.

Link	Document	Ranking
https://www.nytimes.com/2022/02/11/world/middleeast/damiana-nassar-feathers-egypt.html?searchResultPosition=23	<p>She Had Never Acted, and Could Barely Read. Now, She's an Egyptian Movie Star.</p> <p>A stay-at-home mother deprived of an education, Damiana Nassar played a woman much like herself in "Feathers," a Cannes winner that few in Egypt can see.</p>	4

- 6) A rank of 5, although rare, can be awarded when the datapoint is openly disdainful towards the subject and has clear indications of including the author's opinion and preferences rather than reporting the facts. In the below example, the "No Charisma, Please!" has a frustrated tone and claiming that "the two leading candidates are anything but exciting" is extremely insulting and does not have any factual basis (i.e. Germans haven't taken a vote on how exciting the candidates are).

Link	Document	Ranking
https://www.nytimes.com/2021/09/01/world/europe/germany-election-scholz-laschet-baerbock-merkel.html?searchResultPosition=17	<p>It's Election Season in Germany. No Charisma, Please!</p> <p>The race to replace Chancellor Angela Merkel after 16 years in office is the tightest in years. But the two leading candidates are anything but exciting, and that's how Germans like it.</p>	5