

# Relative Movie Attributes

[Link to dataset](#)

[Associated research publication](#)

The dataset consists of sets of movie titles, with each set annotated with a single English soft attribute (subjective descriptive property, such as 'confusing' or 'romantic') and a reference movie. For each set, a crowd worker has placed the movies into three sets: more, equally, and less <attribute> than the reference movie. There are 5,991 such sets, from which one can infer approximately 250,000 pairwise preferences over movies for the 60 distinct soft attributes studied.

<b>PUBLISHER(S)</b> <b>Google LLC</b>	<b>INDUSTRY TYPE</b> <b>Corporate - Tech</b>	<b>INTENDED USE CASE(S), SUITABLE USE CASE(S)</b> This dataset was collected to validate an approach towards scoring soft attributes that could inform improvements in recommendation systems. It is intended for the following use cases: <ul style="list-style-type: none"><li>• Evaluating soft attribute scoring algorithms</li><li>• Training soft attribute scoring algorithms</li><li>• Studying subjectivity of soft attributes</li></ul>
	<b>PRODUCER(S)</b> Filip Radlinski < <a href="mailto:filiprad@google.com">filiprad@google.com</a> > Krisztian Balog Alexandros Karatzoglou	<b>UNSUITABLE USE CASE(S)</b> This dataset is not intended for the following use cases: <ul style="list-style-type: none"><li>• Studying non-English soft attributes</li><li>• Generalizing to soft-attributes outside the domain of popular English-language movies</li><li>• Generating any objective or definitive ranking of movies or anything about them.</li></ul>
	<b>KEY APPLICATION(S)</b> Machine Learning, Preference Elicitation	<b>SOURCE DATASET(S)</b> List of all the datasets that this dataset was derived from: <ul style="list-style-type: none"><li>• <a href="#">Coached Conversational Preference Elicitation</a> (for soft attributes selected)</li><li>• <a href="#">MovieLens 20m</a> (to identify popular movie titles)</li></ul>

<div>PRIMARY DATA FORMAT(S)</div> <div>Textual data</div>	<div>BREAKDOWN</div> <div>Total Instances5,991</div> <div>Number of soft attributes60</div> <div>Distinct movie names300</div>	<div>NATURE OF CONTENT</div> <div>Each instance consists of a set of up to 10 textual movie titles split into three groups, along with a textual soft attribute (e.g. “predictable”), reference movie title and an anonymous rater id.</div>
<div>PRIMARY DATA TYPE(S)</div> <div>Human centered data</div>	<div>EXAMPLE OF ACTUAL DATA POINT*</div> <div><ul style="list-style-type: none"><li>• Rater Id: 100</li><li>• Soft Attribute: romantic</li><li>• Reference Movie: Back to the Future (1985)</li><li>• Less: Jumanji (1995), &lt;other movies&gt;</li><li>• About the same: Apollo 13 (1995), &lt;other movies&gt;</li><li>• More: Titanic (1997), &lt;other movies&gt;</li></ul></div>	<div>HOW TO INTERPRET A DATAPOINT</div> <div>In the example shown, this particular rater considers Jumanji to be less romantic than Back to the Future, Apollo 13 to be similarly romantic to Back to the Future, and Titanic to be more romantic.</div>
<div>DATASET PURPOSE(S)</div> <div>Training</div> <div>Testing</div> <div>Validation</div>	<div>PRIMARY MOTIVATION(S)</div> <div><ul style="list-style-type: none"><li>• Study soft attributes that come up in natural conversational settings.</li><li>• Assess soft attributes as properties that apply to a degree to an item, rather than as absolute properties.</li><li>• Assess subjectivity of soft attributes</li></ul></div>	<div>PRIVACY HANDLING</div> <div><b>Anonymity:</b> Rater ids are integers in the range 1 to 100, not allowing any connection to the actual rater identities.</div>
<div>LICENSE TYPE(S)</div> <div>CC-BY-4.0</div>	<div>VERSION</div> <div>1.0</div> <div>LAST UPDATED</div> <div>11/2020</div> <div>FIRST RELEASE</div> <div>04/2021</div>	<div>LICENSE PERMISSIONS</div> <div><ul style="list-style-type: none"><li>• See <a href="#">CC-BY-4.0 license</a> for license details.</li></ul></div> <div>CHANGELOG (CURRENT VERSION)</div> <div><ul style="list-style-type: none"><li>• Initial version</li></ul></div>

<div>DATA COLLECTION METHOD(S)</div> <div>Crowdsourced</div>		<div>DATA SELECTION CRITERIA</div> <div>Soft attributes were sampled from the Coached Conversational Preference Elicitation dataset, as described in the associated research paper.</div> <div>The 300 movies evaluated were selected as those with most ratings in the MovieLens 20m dataset, as described in the associated research paper. The distribution of these will reflect the distribution of raters of the MovieLens 20m corpus.</div>
<div>SAMPLING METHOD(S)</div> <div>Random sampling</div> <div>Stratified sampling</div>	<div>REPRESENTED POPULATION</div> <div><div>United States93%</div><div>Canada7%</div></div> <div>Gender nor other personal information was not collected.</div>	<div>FILTERING CRITERIA</div> <div>Raters were asked which movies they had seen in an initial filtering step. They were then asked to assess a stratified sample of their seen movies, as described in the associated research paper.</div>
<div>RATING METHOD(S)</div> <div>Non-expert rated</div>	<div>RATER DESCRIPTION(S)</div> <div>Paid crowdworkers</div> <div>RATER TASK(S)</div> <div><div><div>Task 1:</div><div>Raters were presented with a list of movies and asked which they had seen.</div></div><div><div>Task 2:</div><div>Raters were presented with a set of movies, a soft attribute, and a reference movie.</div></div></div>	<div>RATING PROCEDURE - TASK 1</div> <div>Each of 100 raters was shown batches of 20 movies, and asked to mark which of the movies they had seen.</div> <div>RATING PROCEDURE - TASK 2</div> <div>Each of the 100 raters was shown a sequence of tasks. Each task consisted of a set of movies, a soft attribute (“&lt;attribute&gt;”, such as “predictable”), and a reference movie. They were asked to drag-and-drop each movie into one of three buckets, labeled “more &lt;attribute&gt;”, “about the same &lt;attribute&gt;” and “less &lt;attribute&gt;”. Details and illustrations are provided in the associated publication.</div>
<div>VALIDATION METHOD(S)</div> <div>Not validated</div>		<div>VALIDATION POLICY SUMMARY</div> <div>No validation is possible for this task: Raters are giving their own personal opinions of the movies they have seen.</div>

<p>APPLICATION(S)</p> <p><b>Natural language soft attribute interpretation</b></p>	<p>APPLIED METHOD(S)</p> <ul style="list-style-type: none"><li>• Dataset demonstrates practicality of method for collecting assessments of soft attributes.</li><li>• Enables analysis of subjectivity and disagreement among raters for the attributes.</li><li>• Enables further research on acting on user feedback in terms of soft attributes.</li><li>• Ranking model based on weighting over embedding dimensions, as described in detail in associated research publication.</li></ul>	<p>RISKS</p> <p>The publishers have not identified any risks.</p>
--	--	---