

## **Readme: How to use our program?**

Our program can be used to perform the Bechdel-Test, a measure of the representation and portrayal of women in movies. It asks whether a work features at least two women (= first criterion) who talk to each other (= second criterion) about something other than men (= third criterion). We decided to modify the third criterion to whether they are talking to each other for more than three sentences.

The first notebook is the 'Cleaned Dataframe' Notebook, in which we work with the two datasets 'movies\_metadata.csv' and 'credits.csv' which we found on kaggle

(<https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset>). We also work with the Internet Movie Script Database (<https://imsdb.com/>), which contains scripts of 1093 movies. In this notebook, we clean our data in a way that we only have those movies in our dataset for which we also have a script. Here one needs to have the following libraries: pandas, regex, ssl and os.

The second notebook is the 'Preprocessed Dataframe' Notebook, where we work with the previously created cleaned CSVs 'credits\_cleaned' and 'movie\_metadata'. Here we sum up those CSVs in a way that we get one single Dataframe with only the columns that are relevant for us. We are also using the pandas and regex libraries here.

The third notebook is the 'Creating\_Bechdel\_Frame' Notebook where we are working with our cleaned movie data CSV from the previous notebook. In this Notebook, we are conducting the actual Bechdel-Test and creating a final CSV with the results for 750 movies. We need the same libraries as before.

The only notebook that is interesting for the user is the 'plotting\_bechdel' notebook, the last one. For this, one only needs to download our bechdel\_results CSV, as well as the pandas and matplotlib library.

The first part of this notebook one can see several plots of the results of the Bechdel-Test for 750 movies, which our program conducted. The user can see the percentage of how many movies pass the first, second and third criterion of the test in pie charts. Then they can see some regression plots to see whether they might be a correlation between release year and passing of the Bechdel-Test, as well as if there's a correlation of high/low budget and audience voting and the passing of the Bechdel-Test.

The next part is the user input: After the Bechdel-Test has been explained to the user, they can put in a movie, for which they would like to know the result of the Bechdel test. The user is then told which criteria are met and which are not. The user can then also ask which characters exactly speak to each other and in which lines they do so.

## **Programming Journey**

What did we learn throughout the course of the project?

First of all, we learned how immensely difficult it is to work with natural language as data. Analysing the scripts was a great challenge in the sense that there were major stylistic differences between the different film scripts. For example: One movie might include brackets with context information behind the speaking characters while others don't. Then some scripts include the surnames of their characters and some don't. These are just examples to illustrate how the list of particular situations you need to account for just explodes when working with large natural language datasets.

We worked with a lot of nested functions and it was difficult for us to deal with mistakes and error messages in these functions at the beginning. But we learned how to find them via an aimed use of print statements throughout the code. They help identify the exact location of the error and are a great way of controlling how your corrections affect the results.

We also learned how important it is to be aware of the time constraints that go along with analysing large datasets. In our case, some functions took multiple hours to produce a result because they had to go through thousands of data entries. This comes with the additional challenge that the usual trial and error approach of coding gets really time costly. That's why in those cases we learned to really think

through the code beforehand and test it on a few data points for errors before putting it through the whole set.

We realized how important it is to clean one's datasets, especially when working with different datasets. We were working with three different datasets: One with general information about thousands of movies, one with the cast of those movies and one with scripts of about 1000 movies. We decided to clean our data in a way that we only have those data points for which we have both the information about the movie and cast and the corresponding scripts, before the actual conduction of the Bechdel-Test in our program. This way was more time efficient, as one doesn't have to go through the whole dataset every time when running the program.

Although libraries like pandas, matplotlib and regex were touched upon in the lectures, we learned how widely applicable they are and how to really make use of them in a bigger project. For getting the missing genders of some characters in our dataset, we learned how to do web scraping. We were both amazed by how easy it is to extract information from web pages and how many possibilities that opens up.

### **Limitations of our Analysis:**

Our Categorization of the different movies has to be looked at with caution. There are multiple unavoidable sources of potential errors that might distort our results. For example, we concluded that two women are talking with each other if they have consecutive lines in the script's dialog. However, it might be the case that they are just talking after each other in a group setting without actually talking with each other, which would lead to a falsely positive result. On the other hand, we only counted the lines in between two women talking. That way we might've excluded a response of the second woman towards the first woman. This might lead to falsely negative results. However, there was no possibility to distinguish whether the second woman's lines were directed towards the previous speaker or the man following afterwards, without actually semantically analysing the content of the scripts.

Additionally, we could only count the talking characters that were included in our 'credits' dataset. For some movies the cast listing was very extensive and included many characters but some only had few characters listed. That way, there might've been dialog between women, not recognized by our program, as they are not included in our cast-dataframe.

Another point were the missing gender-categorizations of some actresses/ actors. We tried to minimize this error by web scraping from their Wikipedia entries, but as some people didn't have a Wikipedia entry, some missing values remained.

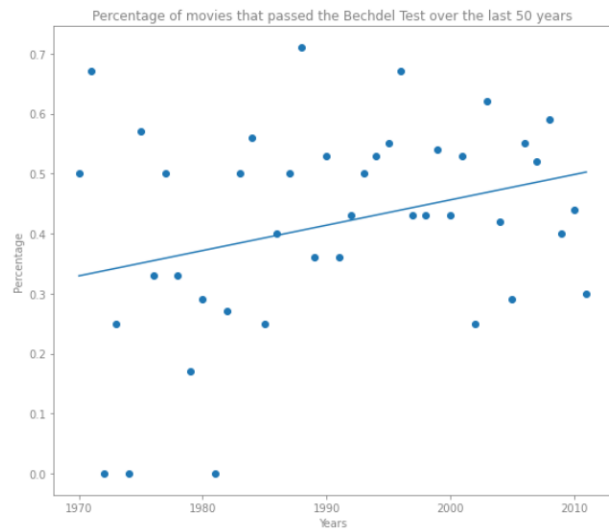
Because our selfly created dataset with the results of the test, which we used for the plotting in the end, is rather small (750 movies), the results are thus not necessarily representative. Another limitation here is, when looking at our plot about how the percentage of movies that passed the Bechdel Test changed over the last 50 years, for example, we might only have few movies from the year 1970, thus those movies have a greater impact on our plot than from a year from which we have many movies in our dataset.

Another thing that could be added to our project would be the programming of the 'original' third criterion. As we changed it from whether women talk to each other about something other than men, to whether they talk to each other for more than three sentences. We had to modify this because otherwise we would have had to semantically analyse the content of the script.

## Our results:

### Most important plots:

Movies with two female characters speaking at least 3 sentences with each other



### User Input:

```
Hello and welcome to our program about the Bechdel Test!
The Bechdel Test is a test to check the representation of women in a movie. It has three criterions:
1. Are there at least two female characters in the movie?
2. Are there female characters talking to each other?
3. Are there female characters talking to each other for more than 3 sentences?
```

```
Results for: Toy Story
```

```
There are/is 0 occurrence(s) in this movie, where women talk with each other.
The movie passes the first criterion: There are at least two female characters in Toy Story
There are no women talking to each other in this movie
There are no women in this movie speaking with each other for more than 3 sentences.
```

```
Hello and welcome to our program about the Bechdel Test!
The Bechdel Test is a test to check the representation of women in a movie. It has three criterions:
1. Are there at least two female characters in the movie?
2. Are there female characters talking to each other?
3. Are there female characters talking to each other for more than 3 sentences?
```

```
Results for: Heat
```

```
There are/is 1 occurrence(s) in this movie, where women talk with each other.
The movie passes the first criterion: There are at least two female characters in Heat
The movie passes the second criterion: There are female characters talking to each other.
The movie passes the third criterion: There are female characters talking to each other for more than 3 sentences.
YAY! Heat passes the whole Bechdel-Test!
['LAUREN (O.S.)\n talks 5 sentence(s) with Justine Hanna from line: 158 till line: 170']
```