

```
1 # === GoodMoives.ipynb ===
2
3 # Dependencie
4 import pandas as pd
5
6
7 # Load in file
8 movie_file = "Resources/movie_scores.csv"
9
10
11 # Read and display the CSV with Pandas
12 movie_file_pd = pd.read_csv(movie_file)
13 movie_file_pd.head()
14
15
16 # List all the columns in the table
17 movie_file_pd.columns
18
19
20 # We only want IMDb data, so create a new table that takes
21 # the Film and all the columns relating to IMDb
22 imdb_table = movie_file_pd[["FILM", "IMDB", "IMDB_norm",
23                             "IMDB_norm_round", "IMDB_user_vote_count"]]
24 imdb_table.head()
25
26
27 # We only like good movies, so find those that scored over 7,
28 # and ignore the norm rating
29 good_movies = movie_file_pd.loc[movie_file_pd["IMDB"] > 7, [
30     "FILM", "IMDB", "IMDB_user_vote_count"]]
31 good_movies.head()
32
33
34 # Find less popular movies--i.e., those with fewer than 20K votes
35 unknown_movies = good_movies.loc[
36     good_movies["IMDB_user_vote_count"] < 20000,
37     ["FILM", "IMDB", "IMDB_user_vote_count"]]
38 unknown_movies.head()
39
40
41 # Finally, export this file to a spread so we can keep track
42 # of out new future watch list without the index
43 unknown_movies.to_excel("output/movieWatchlist.xlsx", index=False)
44
45
```