Page 1/1 Printed for: Amanda Nguyen

45

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