

```
1 # === GoodReadsSummary.py ===
2
3 # Import Dependencies
4 import pandas as pd
5
6
7 #%%
8 # File to Load
9 goodreads_path = "Resources/books_clean.csv"
10
11 # Read the modified GoodReads csv and store into Pandas DataFrame
12 goodreads_df = pd.read_csv(goodreads_path, encoding="utf-8")
13 goodreads_df.head()
14
15
16 #%%
17 # Calculate the number of unique authors in the DataFrame
18 author_count = len(goodreads_df["Authors"].unique())
19
20 # Calculate the earliest/latest year a book was published
21 earliest_year = goodreads_df["Publication Year"].min()
22 latest_year = goodreads_df["Publication Year"].max()
23
24 # Calculate the total reviews for the entire dataset
25 total_reviews = goodreads_df["One Star Reviews"].sum() +
26     goodreads_df["Two Star Reviews"].sum() +
27     goodreads_df["Three Star Reviews"].sum() +
28     goodreads_df["Four Star Reviews"].sum() +
29     goodreads_df["Five Star Reviews"].sum()
30
31
32 #%%
33 # Place all of the data found into a summary DataFrame
34 summary_table = pd.DataFrame({"Total Unique Authors": author_count,
35     "Earliest Year": [earliest_year],
36     "Latest Year": [latest_year],
37     "Total Reviews": [total_reviews]})
38 summary_table
```