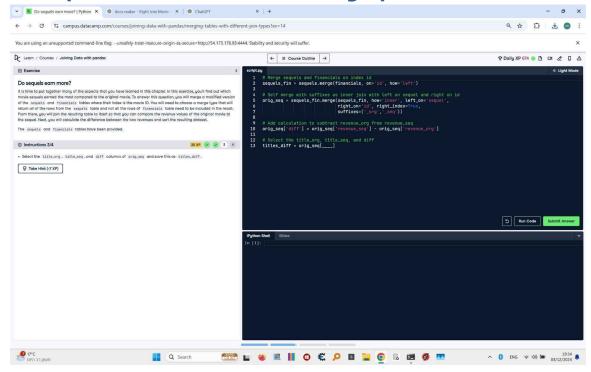
Do Sequels Earn More? - Selecting Specific Columns



Screenshot showing the exercise context for selecting specific columns after merging sequels and financials.

Code Answer:

```
# Merge sequels and financials on index id sequels_fin = sequels.merge(financials, on='id', how='left')
```

Add calculation to subtract revenue_org from revenue_seq orig_seq['diff'] = orig_seq['revenue_seq'] - orig_seq['revenue_org']

```
# Select the title_org, title_seq, and diff columns
titles_diff = orig_seq[['title_org', 'title_seq', 'diff']]
```

Print the resulting DataFrame print(titles diff)

Explanation:

- 1. The first step merges the 'sequels' table with the 'financials' table on the 'id' column using a left join, ensuring all rows from the 'sequels' table are included.
- 2. A self join is then performed on the resulting DataFrame ('sequels_fin') using the 'sequel' column on the left and the 'id' column on the right. The `suffixes=('_org', '_seq')` parameter is used to differentiate columns for the original movie and its sequel.
- 3. The 'diff' column is created by calculating the difference in revenue between the sequel and the original movie.
- 4. Finally, the required columns ('title_org', 'title_seq', and 'diff') are selected and stored in the 'titles_diff' DataFrame. This subset shows the titles of the original movies and their sequels, along with the revenue difference.