

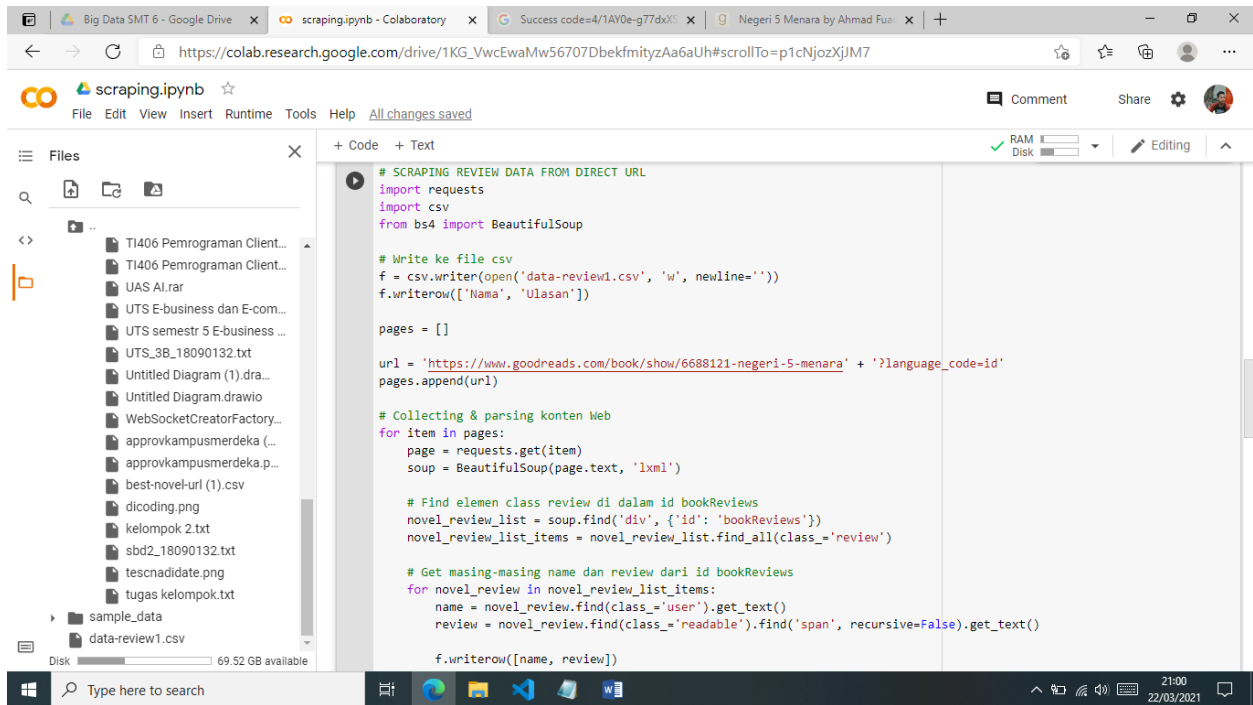
Nama Kelompok : 1. Ananda Mohammad B (18090132)

2. Iqbal Subekhan (18090006)

3. Bagus Fadillah Ahmad (18090008)

4. Nur Sudirman Al Afgani (18090034)

a. Code Scraping URL dan Data



The screenshot displays a Google Colab notebook interface. The left sidebar shows a file explorer with various files, including 'data-review1.csv'. The main area contains a Python script titled '# SCRAPING REVIEW DATA FROM DIRECT URL'. The script imports 'requests', 'csv', and 'BeautifulSoup' from 'bs4'. It defines a function to write data to a CSV file, sets a list of URLs, and then iterates through them to collect and parse book reviews. The script uses 'requests.get()' to fetch pages, 'BeautifulSoup()' to parse the HTML, and 'find()' and 'find_all()' to locate specific review elements. Finally, it writes the extracted names and reviews to the CSV file.

```
# SCRAPING REVIEW DATA FROM DIRECT URL
import requests
import csv
from bs4 import BeautifulSoup

# Write ke file csv
f = csv.writer(open('data-review1.csv', 'w', newline=''))
f.writerow(['Nama', 'Ulasan'])

pages = []

url = 'https://www.goodreads.com/book/show/6688121-negeri-5-menara' + '?language_code=id'
pages.append(url)

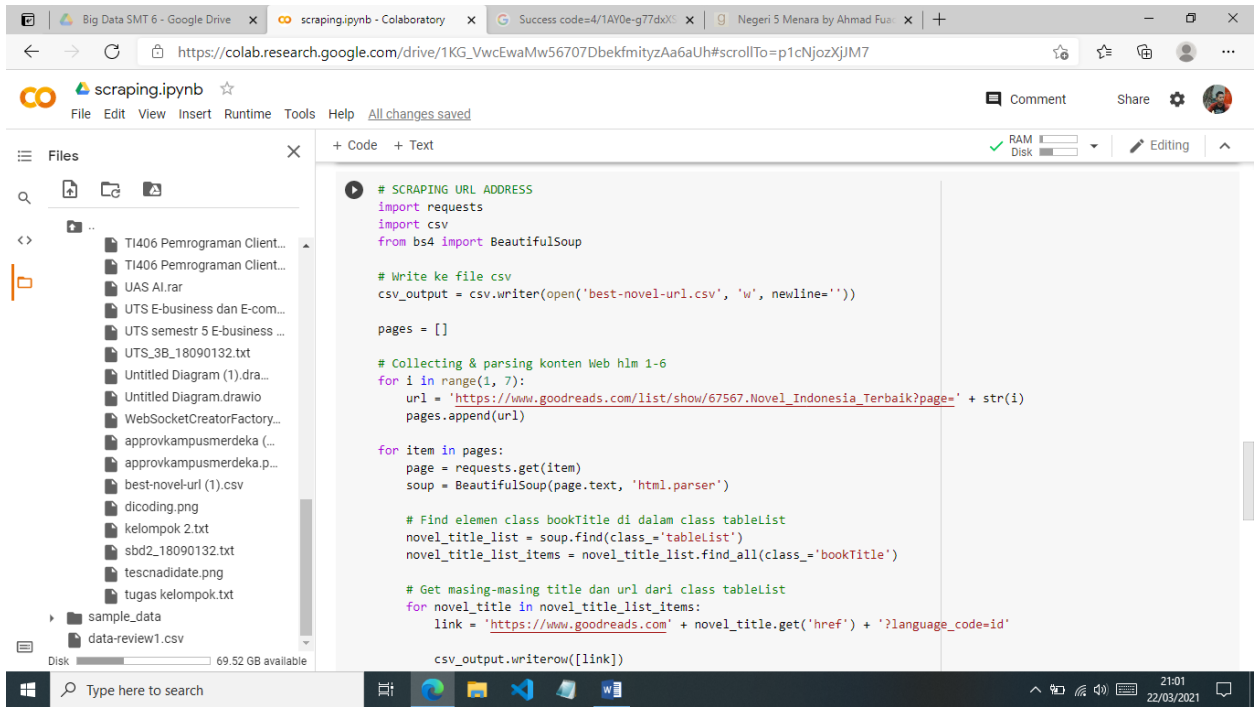
# Collecting & parsing konten Web
for item in pages:
    page = requests.get(item)
    soup = BeautifulSoup(page.text, 'lxml')

    # Find elemen class review di dalam id bookReviews
    novel_review_list = soup.find('div', {'id': 'bookReviews'})
    novel_review_list_items = novel_review_list.find_all(class_='review')

    # Get masing-masing name dan review dari id bookReviews
    for novel_review in novel_review_list_items:
        name = novel_review.find(class_='user').get_text()
        review = novel_review.find(class_='readable').find('span', recursive=False).get_text()

        f.writerow([name, review])
```

b. Hasil Scraping URL



c. Hasil Scraping

