Web data extraction Requirement document

Suggested Code Policy:

1. Copy the Project folder structure from bellow github link.

<https://github.com/hackathonScrapping/BANK.git>

1. Use Python(Mandatory) with BeautifulSoup, Scrapy and Selenium (Feel free to use any other Python libraries)
2. Store Scrapped data with provided file name and CSV format.

Work 1:

**Scarped Data Format (File Name - WellsF\_Data\_Deposit\_<MM\_DD\_YYYY>.csv)**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Date | Bank Name | Product Name | Account Balance | Minimum to Open | Maximum Balance | Rate % | APY % | Bonus Interest Rate | Bonus APY |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

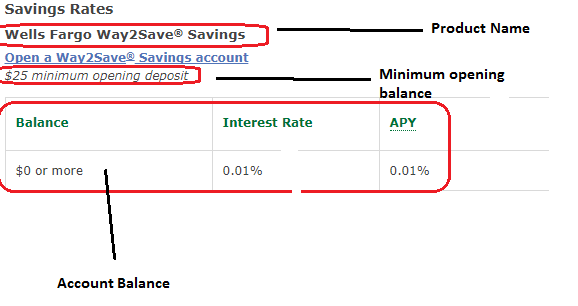
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | |  |  |  |  |  |  | | | |  |
|  |  | |  |  |  |  |  |  | | | |  |
|  | |  | | | | | | | | | |  |
|  |  | |  |  |  | | | |  |  |  |  |

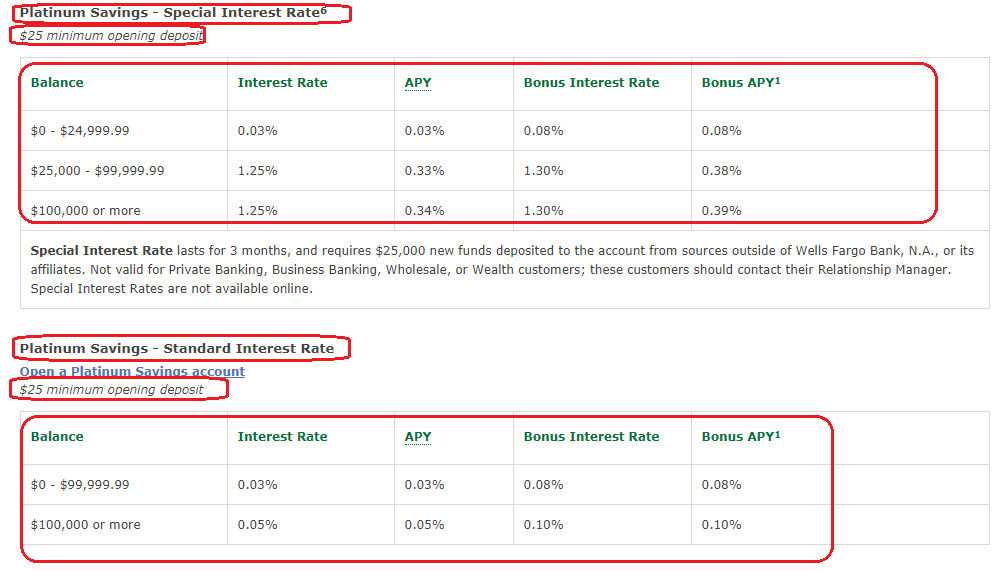
**Steps**

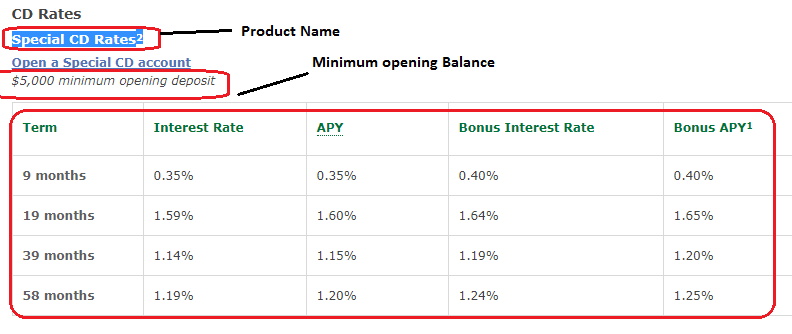
1. Click on below bank site – **PDF document**

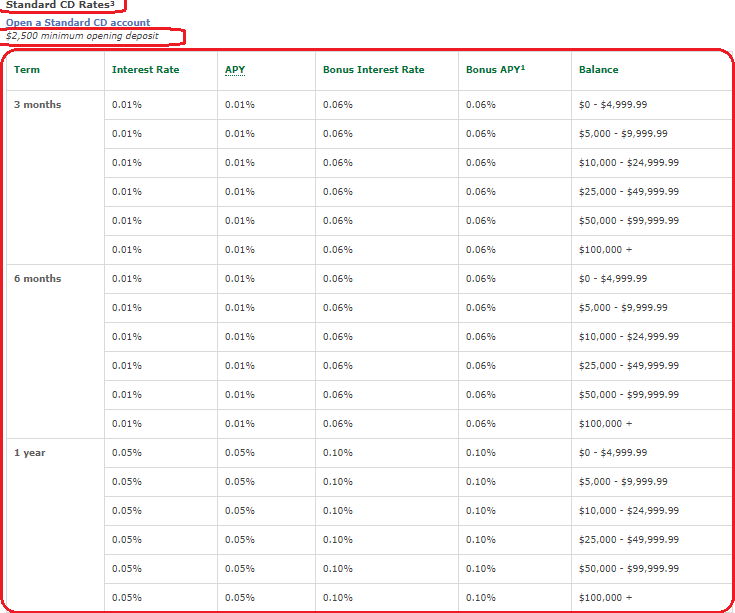
URL: <https://www.wellsfargo.com/savings-cds/rates/>

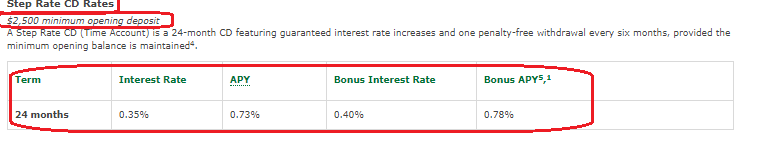
1. Scrape the below highlighted data points







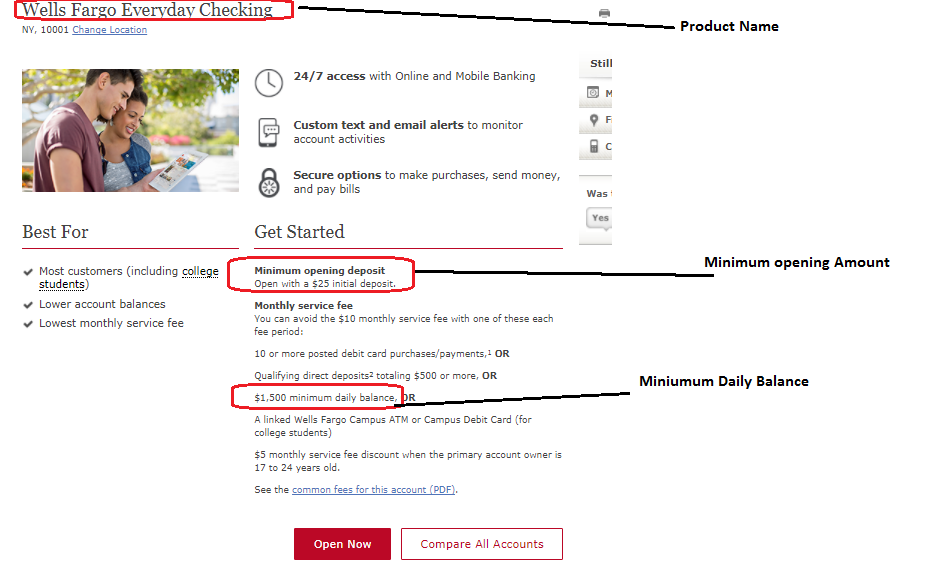




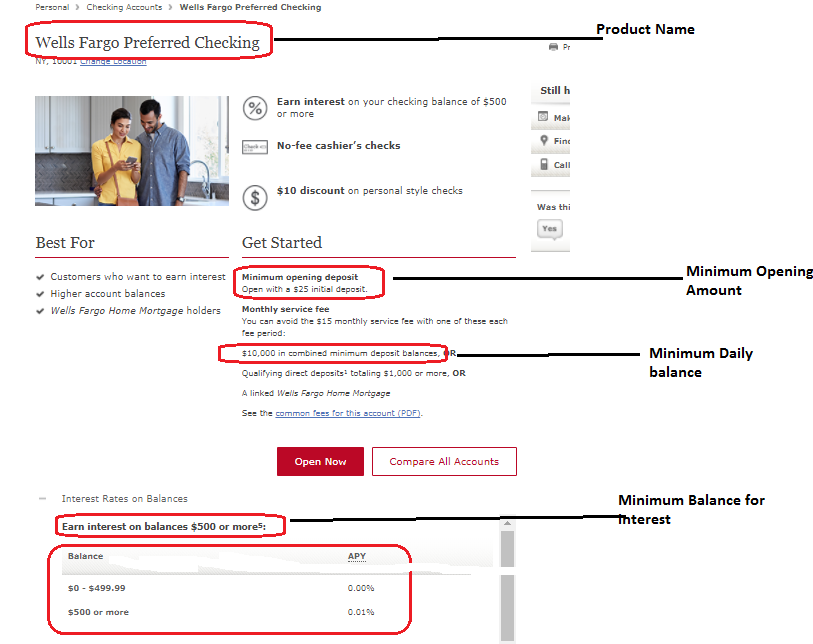
**Checking Accounts**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Date | Bank Name | Interest Type | Repayment Method | Loan amount | State | Loan Term | Credit Score | Loan to Value Ratio (LTV) | Loan Type | Product Name | Interest Rate | APR |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |

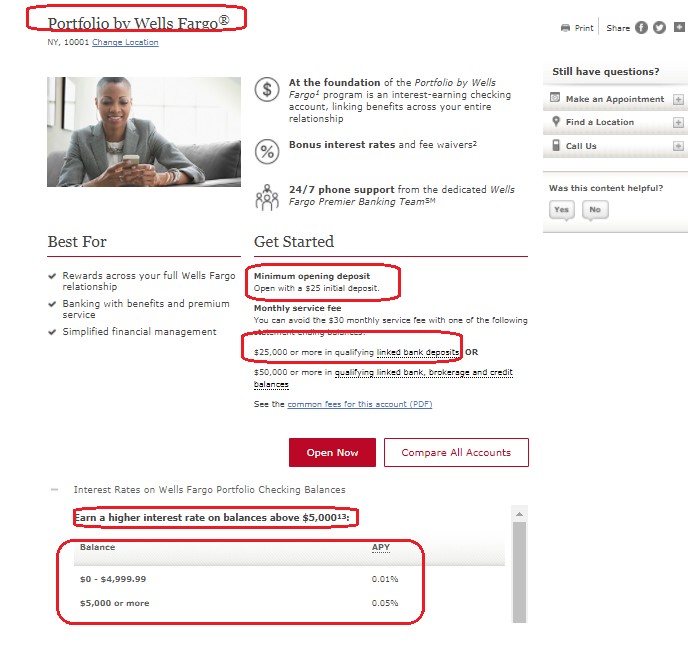
[**https://www.wellsfargo.com/checking/everyday/**](https://www.wellsfargo.com/checking/everyday/)

****

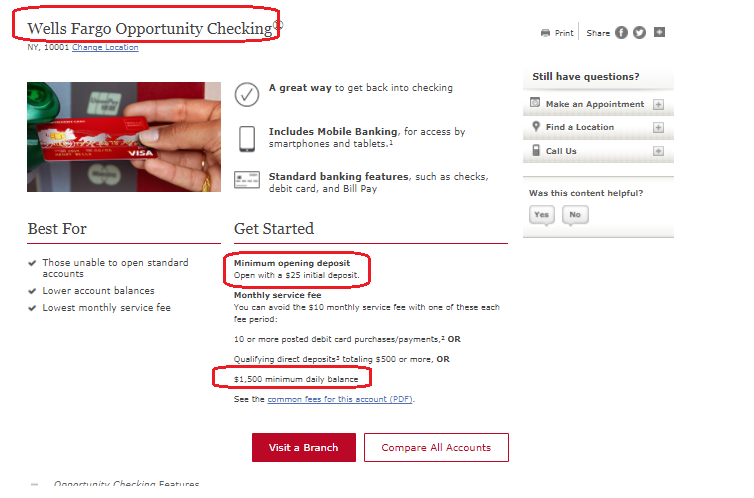
https://www.wellsfargo.com/checking/preferred/



<https://www.wellsfargo.com/checking/portfolio/>



<https://www.wellsfargo.com/checking/opportunity/>



|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Work 2: Capturing Mortgage data  **Capture data for both Mortgage and Refinance along with any other additional data as per below**  **Scarped Data Format (File Name - WellsF\_Data\_Mortgage\_<MM\_DD\_YYYY>.csv)**   |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Date | Bank Name | Interest Type | Repayment Method | Loan amount | State | Loan Term | Credit Score | Loan to Value Ratio (LTV) | Loan Type | Product Name | Interest Rate | APR | |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |   Capture the following:  **Mortgage URL:** <https://www.wellsfargo.com/mortgage/rates/>  **Assumption Loan Amount: - $ 200,000**    **Mortgage URL:** [**https://www.wellsfargo.com/mortgage/rates/purchase-assumptions?prod=5**](https://www.wellsfargo.com/mortgage/rates/purchase-assumptions?prod=5)    **Mortgage URL:** [**https://www.wellsfargo.com/mortgage/rates/purchase-assumptions?prod=6**](https://www.wellsfargo.com/mortgage/rates/purchase-assumptions?prod=6)    **Refinance URL -** [**https://www.bankofamerica.com/mortgage/refinance-rates/**](https://www.bankofamerica.com/mortgage/refinance-rates/)  [**https://www.wellsfargo.com/mortgage/rates/refinance-assumptions?prod=15**](https://www.wellsfargo.com/mortgage/rates/refinance-assumptions?prod=15)  [**https://www.wellsfargo.com/mortgage/rates/refinance-assumptions?prod=16**](https://www.wellsfargo.com/mortgage/rates/refinance-assumptions?prod=16) |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  | |  | |  | |  | |  | |  |  |  |  |  |
|  |  | | |  | |  | |  | |  | |  | | | | | |