

In [3]:

```

# IMPORT ALL MODULES
import pandas as pd
import requests
from bs4 import BeautifulSoup
import selenium
from selenium import webdriver
import time

# DEFINE FUNCTION TO GENERATING URL FOR WHOLE PAGE
def get_url(productsearch):
    productsearch = productsearch.replace(' ', '%20')
    productsearch = productsearch.replace('"', '%27')
    template = 'https://shopee.com.my/search?keyword={}'
    url = template.format(productsearch)
    return url

def get_wholepg(card):

    # PRODUCT IMG INFO
    pImg = card.find('img')
    product_image = pImg['src']

    # PRODUCT NAME INFO
    product_name = card.find('div', '_10Wbs- _5SSWfi UjjMrh').text.strip()

    # PRODUCT PRICE
    product_price = card.find('div', 'zp9xm9 xSxKlK _1heB4J').text.strip()

    # THE LINK FOR EACH CARD
    product_url = card.a.get('href')
    prodUrl = 'https://shopee.com.my'+product_url

    # SAVE IT ALL INTO A DATAFRAME
    product_info = (product_image, product_name, product_price, prodUrl)

    return product_info

# DEFINE THE MAIN FUNCTION (SEARCH INPUT)
def main(productsearch):
    records = [] # EMPTY LIST FOR DTFRAME
    url = get_url(productsearch)

    # NEED SELENIUM TO GENERATE THE WORKING LINK
    driver = webdriver.Chrome(executable_path='chromedriver.exe')
    driver.get(url)
    driver.maximize_window()
    # WAIT 5 SECS
    time.sleep(5)
    # CLICK ENG BUTTON AUTOMATICALLY
    btn = driver.find_element_by_xpath('//*[@id="modal"]/div[1]/div[1]/div/div[3]/div[1]')
    btn.click()

    # DEFINE FUNCTION TO CALL CARDS INFORMATION
    soup = BeautifulSoup(driver.page_source, 'html.parser')
    product_cards = soup.find_all('div', 'col-xs-2-4')

    for everyCard in get_wholepg:
        productDetails = get_wholepg(everyCard)
        records.append(productDetails)

```

```
# USING PANDAS DATAFRAME SAVE THE INFO IN CSV FILE
col = ['Product_Image', 'Product_Name', 'Product_Price', 'OfficialSite_Link']

shopee_data = pd.DataFrame(records, columns=col)

shopee_data.to_csv('C:\\Users\\Apeh\\Documents\\csv_savefiles', encoding='utf-8')
```