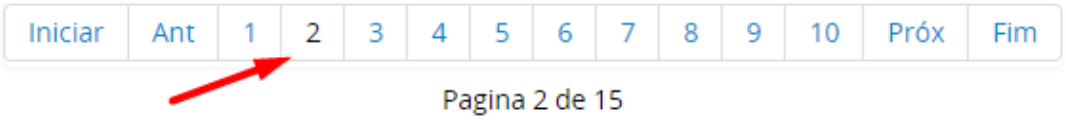
PART I: Data scraping

Form the following link:

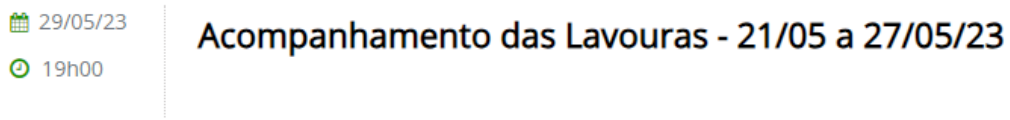
https://www.conab.gov.br/info-agro/safras/progresso-de-safra

You will find weekly publications for a website called CONAB. In each one, we are interested in the planting pace for Corn 1 (Milho 1) only:

Start from the second page:

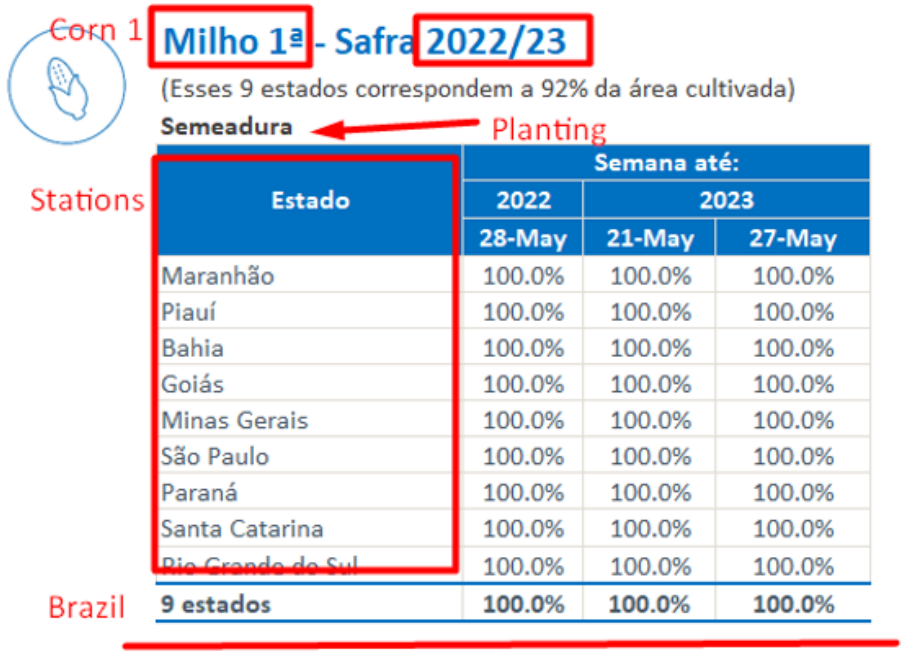


The publication on the 29th of May 2023:



1- **find the links that start with Plantio e Colheita:**

2- **An Excel file will be downloaded. (When you write the script don't download the file but use a buffer)**
3- **Open the Excel file and retrieve the following information: (use python to retrieve the data and not manually)**



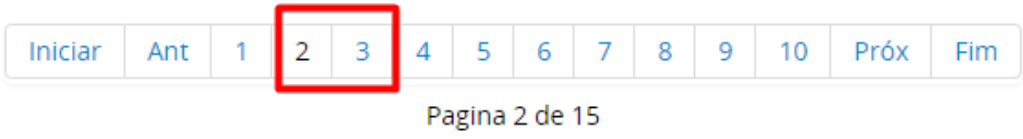
Estado	Semana até:		
	2022	2023	
	28-May	21-May	27-May
Maranhão	100.0%	100.0%	100.0%
Piauí	100.0%	100.0%	100.0%
Bahia	100.0%	100.0%	100.0%
Goiás	100.0%	100.0%	100.0%
Minas Gerais	100.0%	100.0%	100.0%
São Paulo	100.0%	100.0%	100.0%
Paraná	100.0%	100.0%	100.0%
Santa Catarina	100.0%	100.0%	100.0%
Rio Grande do Sul	100.0%	100.0%	100.0%
9 estados	**100.0%**	**100.0%**	**100.0%**

4- For each table you need to have the following structure:

Date	Location	Crop	Value
27-05-2023	Maranhao	2022/23	1
27-05-2023	Piaui	2022/23	1
27-05-2023	Bahia	2022/23	1
27-05-2023	Goias	2022/23	1
27-05-2023	Minas Gerais	2022/23	1
27-05-2023	Sao Paulo	2022/23	1
27-05-2023	Parana	2022/23	1
27-05-2023	Santa Catarina	2022/23	1
27-05-2023	Rio Grande do Sul	2022/23	1
27-05-2023	Brazil	2022/23	1

Just take the last column which is the last date.

5- Try to run the script (dynamically) on the 2nd and 3rd pages and append every result you have to have an Excel file that has all the values.



PS:

- **You cannot use Selenium.**
- **You find attached the expected output; if you couldn't finish the scraping part you can use it to finish the Power BI task.**

PART II: Power Bi

You will find a dataset that contains these data with more historical.

Use it in this task. The file is named: "data.xlsx"

On the first page, you have to build a table that has the following structure:

Corn 1 Harvesting for crop 2022/23 published at 27 May 2023 @DNEXT

Product	Corn 1				
Location	Current	Last week	WoW	Average	Vs average
Brazil	82%	77%	5% ↑	85%	-3% ↓
Bahia	73%	71%	2% ↑	79%	-6% ↓
Goiás	93%	50%	43% ↑	97%	-4% ↓
Maranhão	34%	25%	9% ↑	47%	-13% ↓
Minas Gerais	99%	97%	2% ↑	100%	-1% ↓
Parana	98%	96%	2% ↑	100%	-2% ↓
Piauí	35%	30%	5% ↑	33%	2% ↑
Rio Grande do Sul	93%	90%	3% ↑	92%	1% ↑
Santa Catarina	100%	100%	0% →	100%	0% →
São Paulo	100%	100%	0% →	100%	0% →

Go to the line chart view

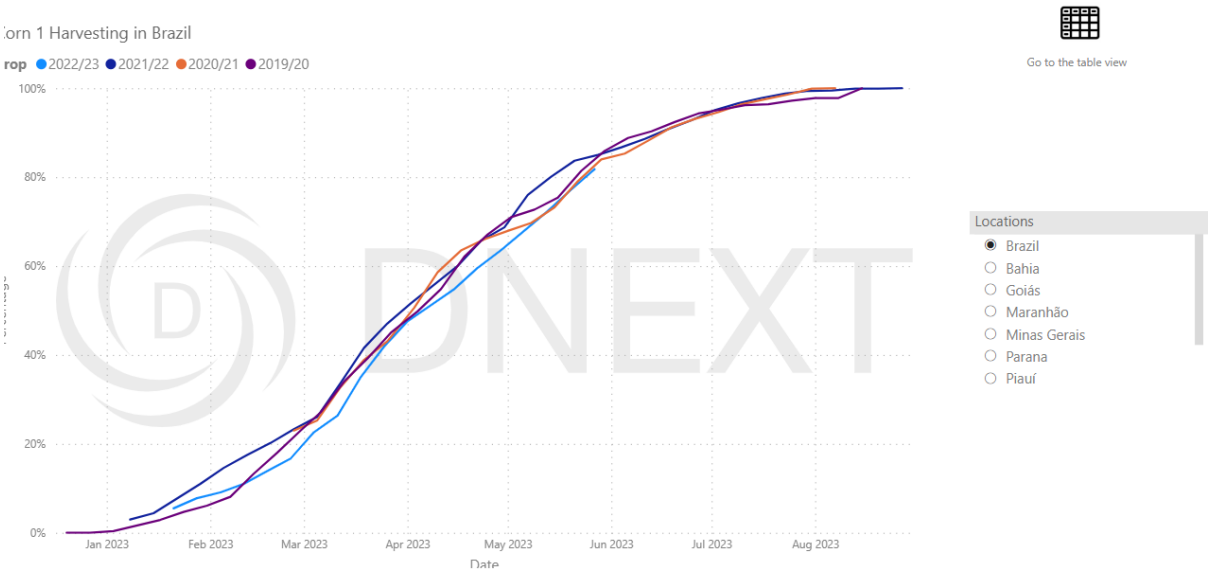
Published date

- (●) 27 May 2023
- ( ) 20 May 2023
- ( ) 13 May 2023
- ( ) 6 May 2023
- ( ) 29 April 2023
- ( ) 22 April 2023
- ( ) 15 April 2023
- ( ) 8 April 2023
- ( ) 1 April 2023

When the user selects a published date:

- **The current column displays the values of that selected week.**
- **Last week displays the values of the previous week according to the selection.**
- **WoW is Week over Week it's Current – Last Week**
- **The Average column displays the average value on the week selected in the previous years (you need to use the week number function in Dax, it's ok if you didn't manage to do it)**
- **Vs average is current – Average.**

When you click on this icon  **you should be directed to the following graph:**



The graph should display a seasonal chart for the different years.

Please send your submission as follows:

- **Respond to the email we sent you with the test.**
- **Attach a zip/rare file: that contains the python script file + its output in the name (output.xlsx) + the pbix file for the dashboard.**
- **Your submission should be sent in maximum 3h after the receiving the test.**

Good Luck