



Sports Data Analysis

Week 2





Good Sources

- <https://www.sports-reference.com/> (by far the best)
 - [fangraphs.com](https://www.fangraphs.com) (just baseball)
 - ESPN API
 - Pro league APIs
- 
- 

Sports Reference Tip

For each table, click on “Share & Export” at the top and then “Get table as CSV (for Excel)”

From here, copy and paste the table into google sheets and then download it as a CSV

Game averages

Att	Yds	Avg
8.0	253.5	6.7
3.0	77.5	2.3
5.0	+176.0	+4.4

Share & Export

- Modify, Export & Share Table
- Get as Excel Workbook
- Get table as CSV (for Excel)
- Get Link to Table
- About Sharing Tools
- Video: SR Sharing Tools & How-to
- Video: Stats Table Tips & Tricks
- Data Usage Terms

Export

%	Yds	TD
6.9	733	2
0.0	0	0
6.3	733	2

Turnovers

Att	Int	Tot
0.5	0.3	0.8
0.5	1.5	2.0
0.0	-1.2	-1.2

Sports Reference Cntd.

Team Stats

Most values are per game averages

Share & Export ▾

[Glossary](#)

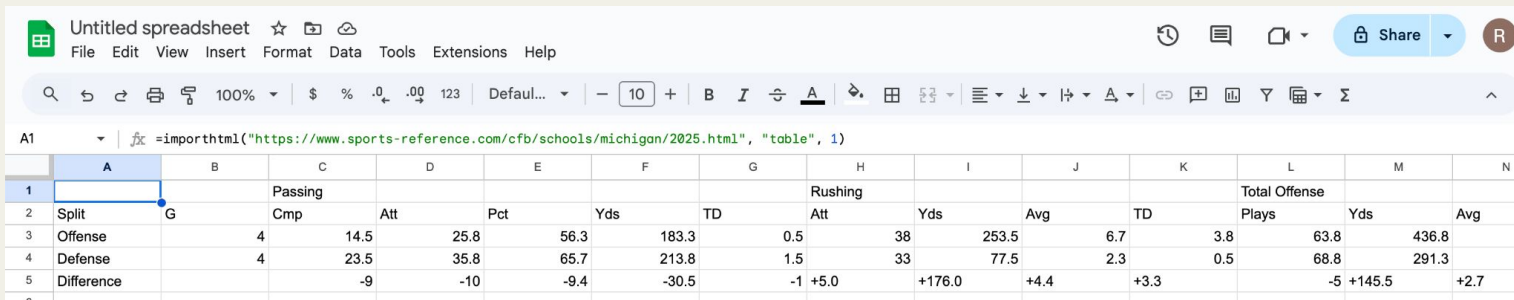
[Reload](#) page to return to the table-formatted data.

-- When using SR data, please cite us and provide a link and/or a mention.

```
,Passing,Passing,Passing,Passing,Passing,Rushing,Rushing,Rushing,Rushing>Total Offense>Total Offense>Total Offense,First Downs,First Downs,First Downs,First Downs,Per  
plit,G,Cmp,Att,Pct,Yds,TD,Att,Yds,Avg,TD,Plays,Yds,Avg,Pass,Rush,Pen,Tot,No.,Yds,Fum,Int,Tot  
ffense,4,14.5,25.8,56.3,183.3,0.5,38.0,253.5,6.7,3.8,63.8,436.8,6.9,9.0,10.8,1.0,20.8,4.8,45.3,0.5,0.3,0.8  
ffense,4,23.5,35.8,65.7,213.8,1.5,33.0,77.5,2.3,0.5,68.8,291.3,4.2,10.5,5.5,2.5,18.5,2.5,17.0,0.5,1.5,2.0  
ifference,,-9.0,-10.0,-9.4,-30.5,-1.0,+5.0,+176.0,+4.4,+3.3,-5.0,+145.5,+2.7,-1.5,+5.3,-1.5,+2.3,+2.3,+28.3,0.0,-1.2,-1.2
```

In the example from the previous slide, you would get this output and would copy the table underneath the "-- When using SR data.." into google sheets for saving as CSV

Other Scrapping Tips



Untitled spreadsheet

File Edit View Insert Format Data Tools Extensions Help

100% 123 Default... 10 B I A

A1 `=importhtml("https://www.sports-reference.com/cfb/schools/michigan/2025.html", "table", 1)`

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1			Passing					Rushing				Total Offense		
2	Split	G	Cmp	Att	Pct	Yds	TD	Att	Yds	Avg	TD	Plays	Yds	Avg
3	Offense		4	14.5	25.8	56.3	183.3	0.5	38	253.5	6.7	3.8	63.8	436.8
4	Defense		4	23.5	35.8	65.7	213.8	1.5	33	77.5	2.3	0.5	68.8	291.3
5	Difference			-9	-10	-9.4	-30.5	-1	+5.0	+176.0	+4.4	+3.3	-5	+145.5
6														+2.7

In google sheets, use the function IMPORTHTML (seen above), with the usage:



`=importhtml("url", "table", index)` where url must be in quotes and index is (I believe from my testing) the table you want in order top to bottom starting from 1



Another (harder) Way to Scrape

There is a Python library called BeautifulSoup that is built for webscraping data!

I have code that uses it in my github repository under each dataset's folder that you can use to model your own code after (but try the other options first!!):
<https://github.com/MichiganDataScienceTeam/F25-SportsDataAnalysis/tree/main/Data>



Questions? Email me!



CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)

Please keep this slide for attribution

