Website: [Box Office Mojo](https://www.boxofficemojo.com/)

Questions:

1. What holiday had the highest or lowest cumulative gross in 2022?
2. Which month had the highest or lowest cumulative gross?
3. What year had the highest or lowest average gross?
4. What week in 2022 had the highest or lowest overall gross?
5. Which movie had the most days as the #1 release in 2022?

Plan:

[Yearly top gross](https://www.boxofficemojo.com/year/?ref_=bo_nb_hm_secondarytab)

[2022 Box Office](https://www.boxofficemojo.com/year/2022/?ref_=bo_yl_table_2)

[2022 Monthly Box Office](https://www.boxofficemojo.com/month/by-year/2022/?grossesOption=calendarGrosses)

[2022 Weekly Box Office](https://www.boxofficemojo.com/weekly/by-year/2022/)

[2022 Holiday Box Office](https://www.boxofficemojo.com/holiday/by-year/2022/?grossesOption=calendarGrosses)

id=”table”

class\_='imdb-scroll-table'

Pull data example:

url='https://www.boxofficemojo.com/daily/2022/?view=year'

    req=requests.get(url)

    content=req.text

    soup=bs(content)

    rows=soup.find\_all('tr')

    daily\_movies = []

    for row in rows:

        data\_row = {}

        data = row.find\_all('td')

        if len(data) == 0:

            continue

        data\_row['date'] = data[0].text

        data\_row['day'] = data[1].text

        data\_row['top10gross'] = data[3].text.split('$')[1]

        data\_row['releases'] = data[6].text

        data\_row['top\_release'] = data[7].text

        data\_row['gross'] = data[8].text.split('$')[1]

        daily\_movies.append(data\_row)

    daily\_df = pd.DataFrame(daily\_movies, columns = ['date', 'day', 'top10gross', 'releases', 'top\_release', 'gross'])

    return daily\_df

<https://www.boxofficemojo.com/robots.txt>

Box Office Mojo allows all.