

SteelSpotter

Lorry Zou (lorryz), Xiaoyu Song (xiaoyuso), Ruofeng Xu (ruofengx), Yuchen Ma (yuchenm)

Our app, SteelSpotter, provides ticket prices and games information for all Pittsburgh sports fans. We scraped the data of Panthers men's basketball games and Steelers' football games from SeatGeek and StubHub. All the data displayed in our application were pre-scraped on Dec 4th, 2024. However, users can still access the latest game and ticket information through online web-scraping and download the data as files by themselves. Additionally, SteelSpotter provides features for users to browse local hotels and explore statistics of their favorite teams and players, creating a seamless and informative experience for sports enthusiasts.

Link to our demonstration video: <https://youtu.be/e6wfftYsw1I>

Installation

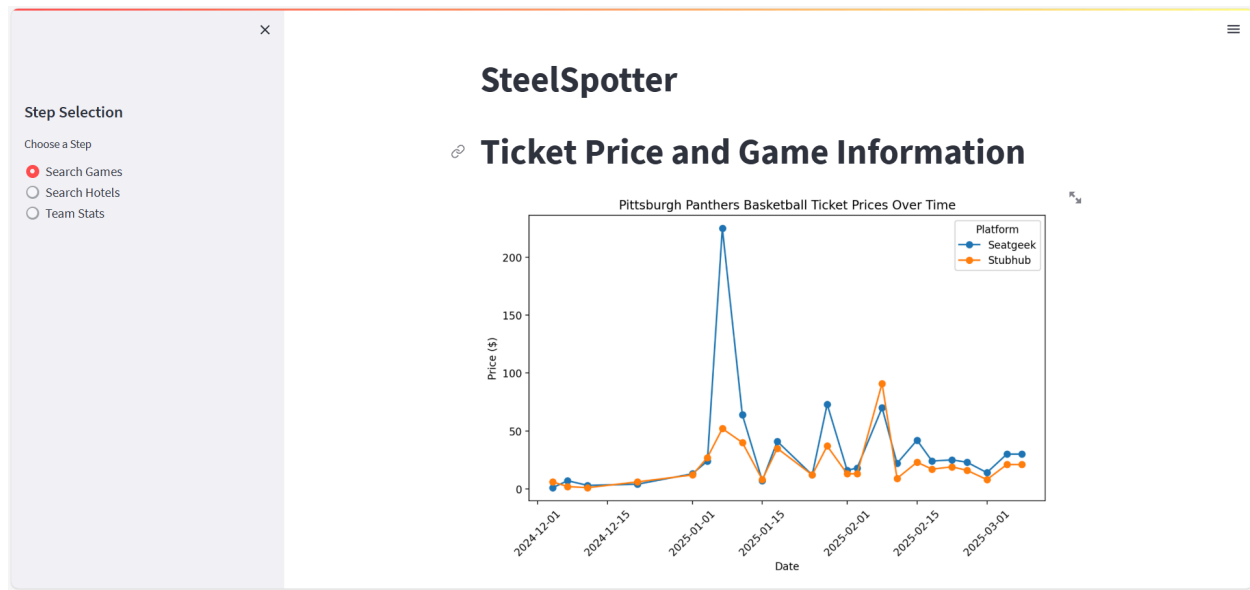
Before running our application, you should install the following Python packages:

- Pandas: `pip install pandas`
- Matplotlib: `pip install matplotlib`
- Playwright: `pip install playwright` - AND THEN RUN - `playwright install`
- Streamlit: `pip install streamlit`
- Selenium: `pip install selenium`
- Altair: `pip install altair==4`
- BeautifulSoup: `pip install beautifulsoup4`

Run

Our app is developed with Streamlit. After installing all the required packages above, you can unzip our submission and run the `main_file.py` by running `streamlit run main_file.py` in the terminal, and then our application will be running on your local browser. If it asks you for your email address, just press "Enter" to skip it.

You'll see the home page once you have SteelSpotter running:



To close the application, simply shut down the browser tab.

Ticket Scraper and Data Processor

Note: You don't need to run the scraping scripts as the scraped data has already been provided in the ZIP file. The instructions below describe the data collection process for your reference.

SeatGeek

SeatGeek has strict bot detection mechanisms, and if you scrape too often, it may flag your activity as robotic. To avoid this, follow the steps below:

```
python seatgeek_panthers_scrape.py
```

After completing the basketball scrape, clear your browser cookies and wait a few minutes to reset your session. Then run:

```
python seatgeek_steelers_scrape.py
```

Stubhub

Same for stubhub, but stubhub won't recognize us as robots. So the scraper can be done in `main_file`. But it takes a lot of time.

Data Processor

As long as you get basketball and football tickets data from 2 platform, you can run the following commands to process and combine all the tickets data:

```
python combine_tickets_data.py
```