

# Project aim and outline

---

**Aim: To learn more about web scraping.**

**Outline:**

- > Scrape game titles from IGN website.**
- > Scrape game data from Steam platform.**
- > Explore data using jupyter notebook and Pandas.**

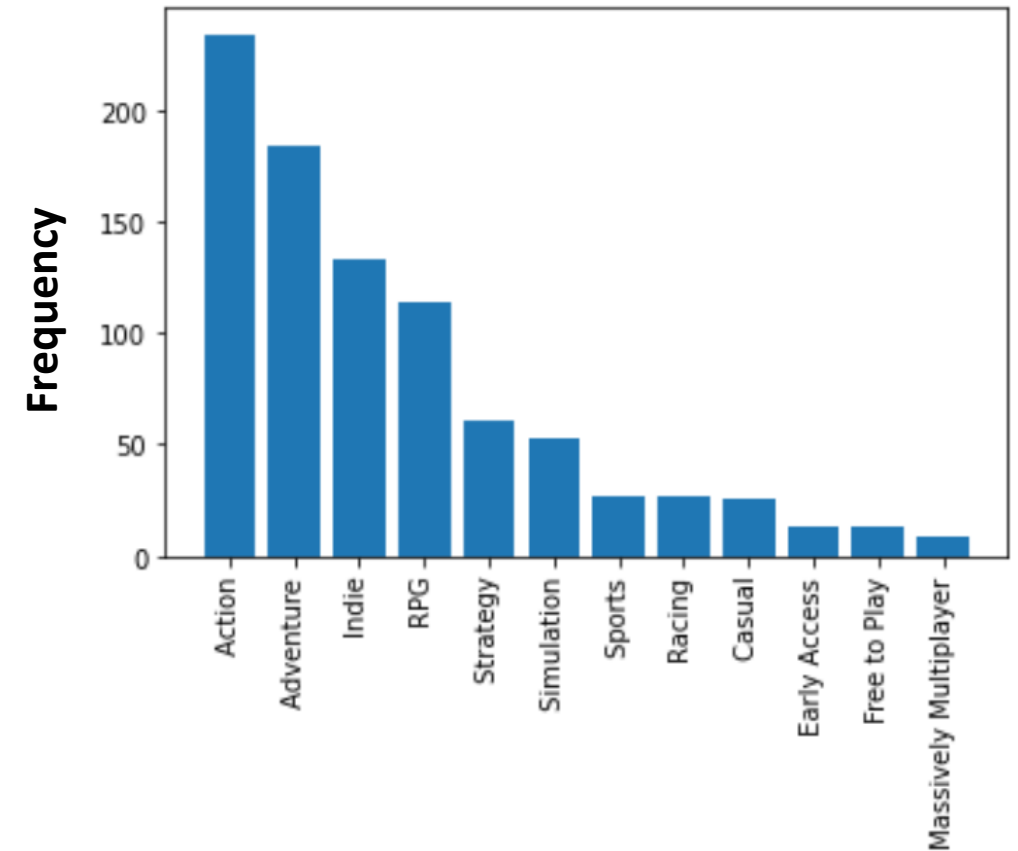
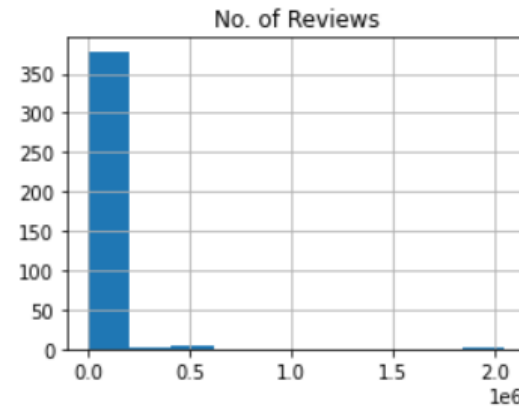
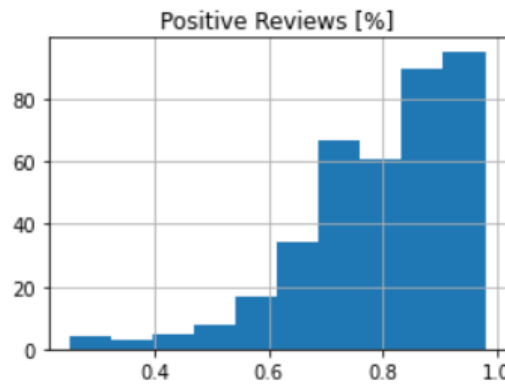
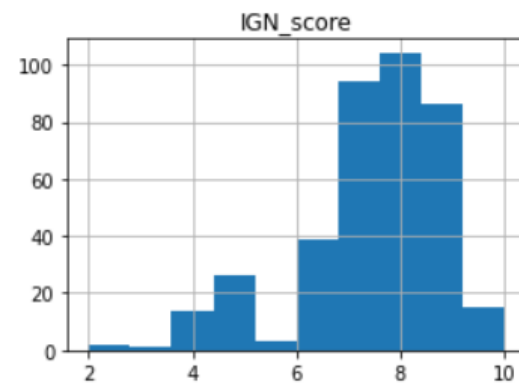
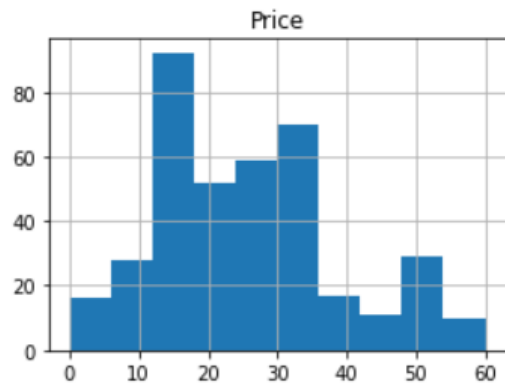
# Web scraping with Selenium

---

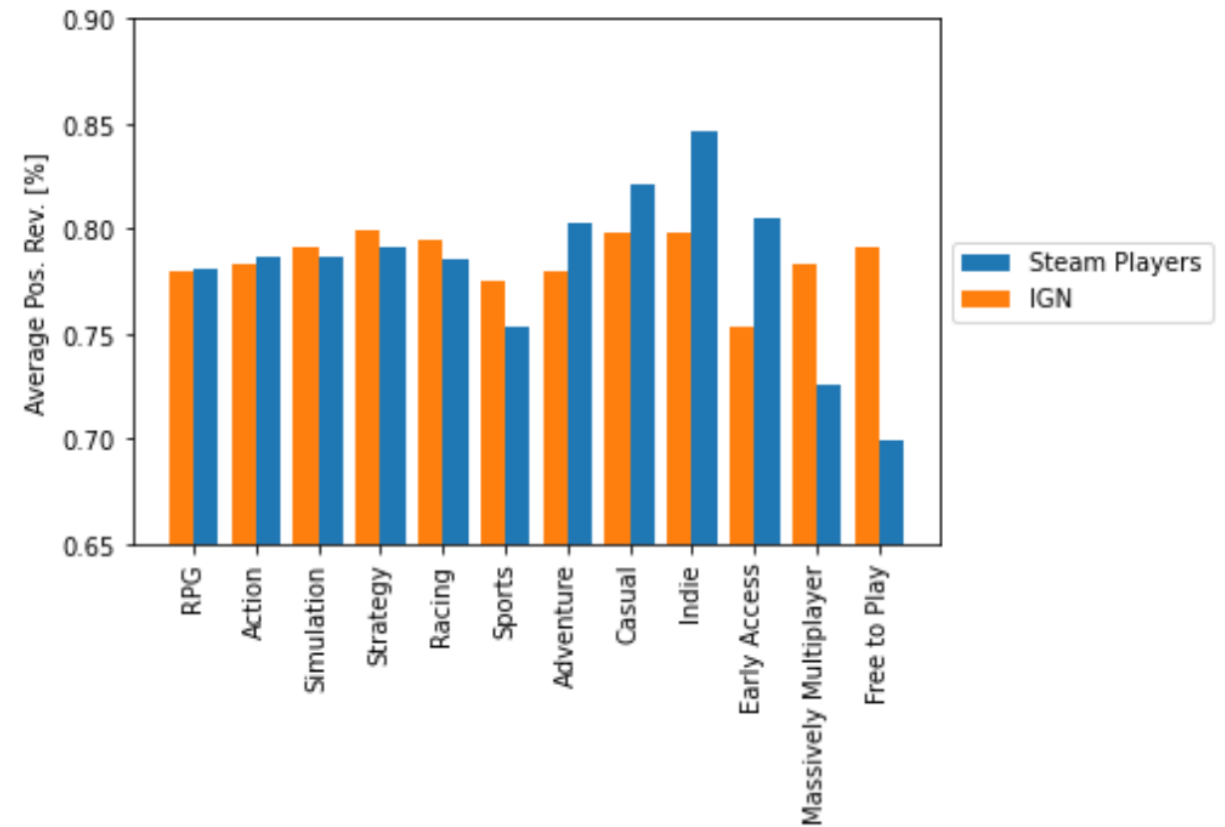
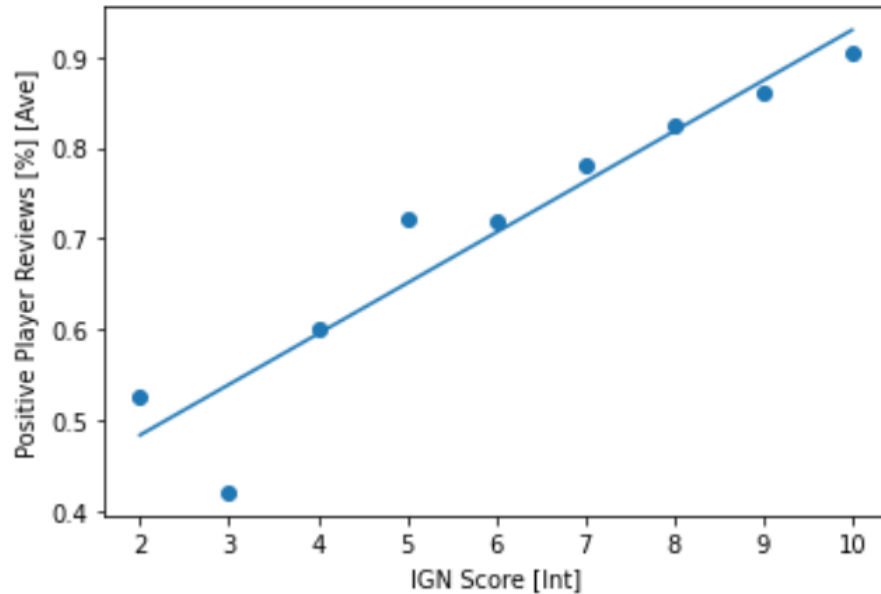
**Selenium is web browser automation tool.**

- 1. Open IGN webpage and scroll. Scrape game titles.**
- 2. Clean broken game tiles.**
- 3. Open Steam & scrape game data from titles.**
- 4. Clean data in jupyter notebook and pandas.**

# Exploring game data: Distributions



# Exploring game data: Players VS IGN



# Conclusions overview

---

- > Selenium can be used to scrape data from webpages with dynamic elements.
- > It is very helpful to know what your using the data for before you scrape it.
- > Scaped data will require lots of cleaning. Manual cleaning can take up a lot of time.
- > Keep as much data as possible. Likely I could have kept some of the dropped data.

---

**Please see jupyter notebooks for detailed explanation of how data was cleaned, scraped and, explored.**

**This was a brief summary.**