
Potential Ebay Scams

Zion Calvo (zcalvo@ucsc.edu)

Mark Mileyev (mmileyev@gmail.com)

Questions:

- How prevalent are fraudulent listings on ebay?

- Which categories are more prone to having scams?

- Are there any patterns we could find in these items?

Data Sources:

- Ebay API: <https://developer.ebay.com/DevZone/finding/CallRef/findCompletedItems.html>

- Scraped Descriptions: [https://vi.vipr.ebaydesc.com/ws/eBayISAPI.dll?ViewItemDescV4&item=\(ITEM#\)](https://vi.vipr.ebaydesc.com/ws/eBayISAPI.dll?ViewItemDescV4&item=(ITEM#))

- Scraped Seller History: [https://feedback.ebay.com/ws/eBayISAPI.dll?ViewFeedback2&userid=%27%20+\(USER\)\)](https://feedback.ebay.com/ws/eBayISAPI.dll?ViewFeedback2&userid=%27%20+(USER)))

Code / Viz References

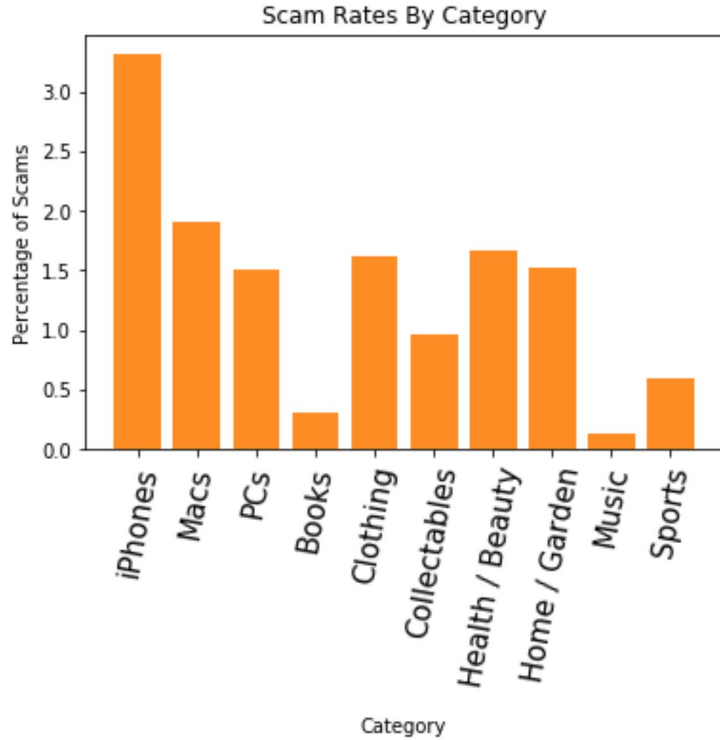
[Kaggle](#)

- Shows Auction Data not Buy it now items
- Gave data structure inspiration

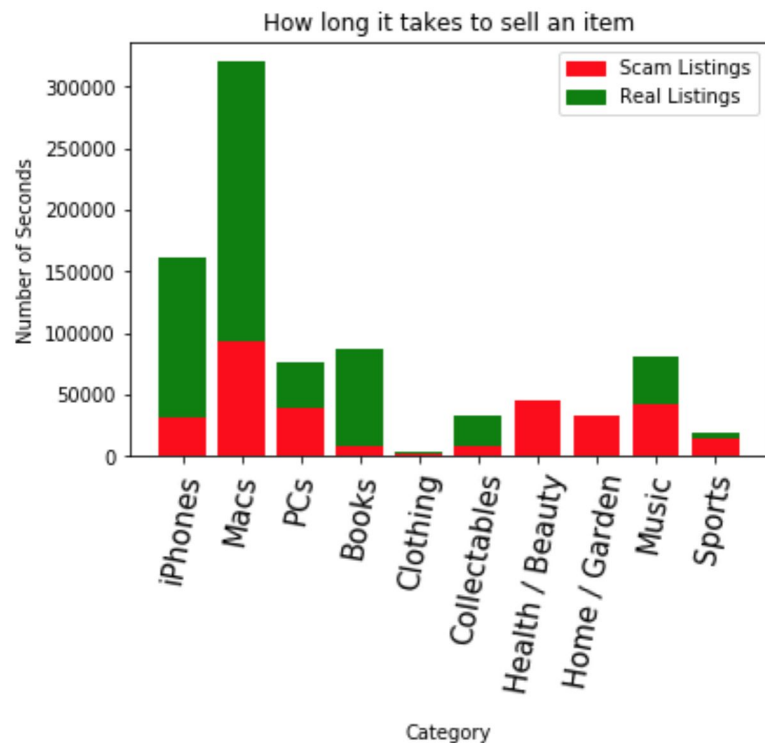
[Jay Grossman](#)

- Used this machine learning project for inspiration on how to structure our scraped data CSV's
 - Potentially useful graphs
-

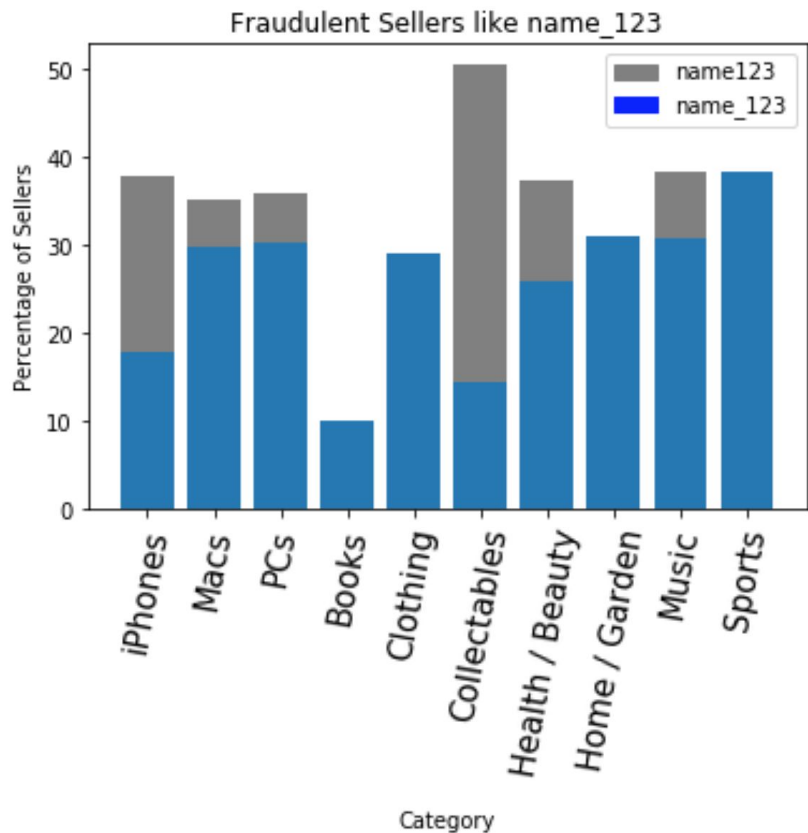
RESULTS



- iPhones had the highest potential scam rate at 3.3%
- Books and Music categories appear to have very little fraudulent activity
- Other categories vary in the 1-2% range



- Potentially Fraudulent iPhone listings sell 81% faster than other iPhone listings
- Home and Garden categories take longer for potentially fraudulent items to sell
- Overall, legitimate items take longer due to higher prices



- Many of the fraudulent seller names follow a pattern suggesting they may be computer generated
- About 1 in 3 scam listings is posted by a seller following this naming pattern
- This naming pattern appears less frequently in legitimate listings

Libraries Used

Pandas

Requests

BeautifulSoup

Matplotlib

SKLearn

Accomplishments

Data Wrangling

Converging multiple data sources into one CSV for each ebay category

Analysis

Using Pandas to effectively answer our questions and filter relevant items

Visualization

Plotted our findings in a format that is easy to understand and visualize

Challenges

Scraping / Building our datasets

Building multiple CSV's through a multistep process with unpredictable data was a new problem for us

Each scraped item went through a process of first being pulled from ebay API, storing attributes from the API call, then using beautifulsoup to look up the description page of the item and adding onto the attributes. After that, the seller id for the item gets their last 5 reviews scraped with beautifulsoup and requests
