NBA Shot Data Visualization

HACK PRINCETON WORKSHOP 4/2/16
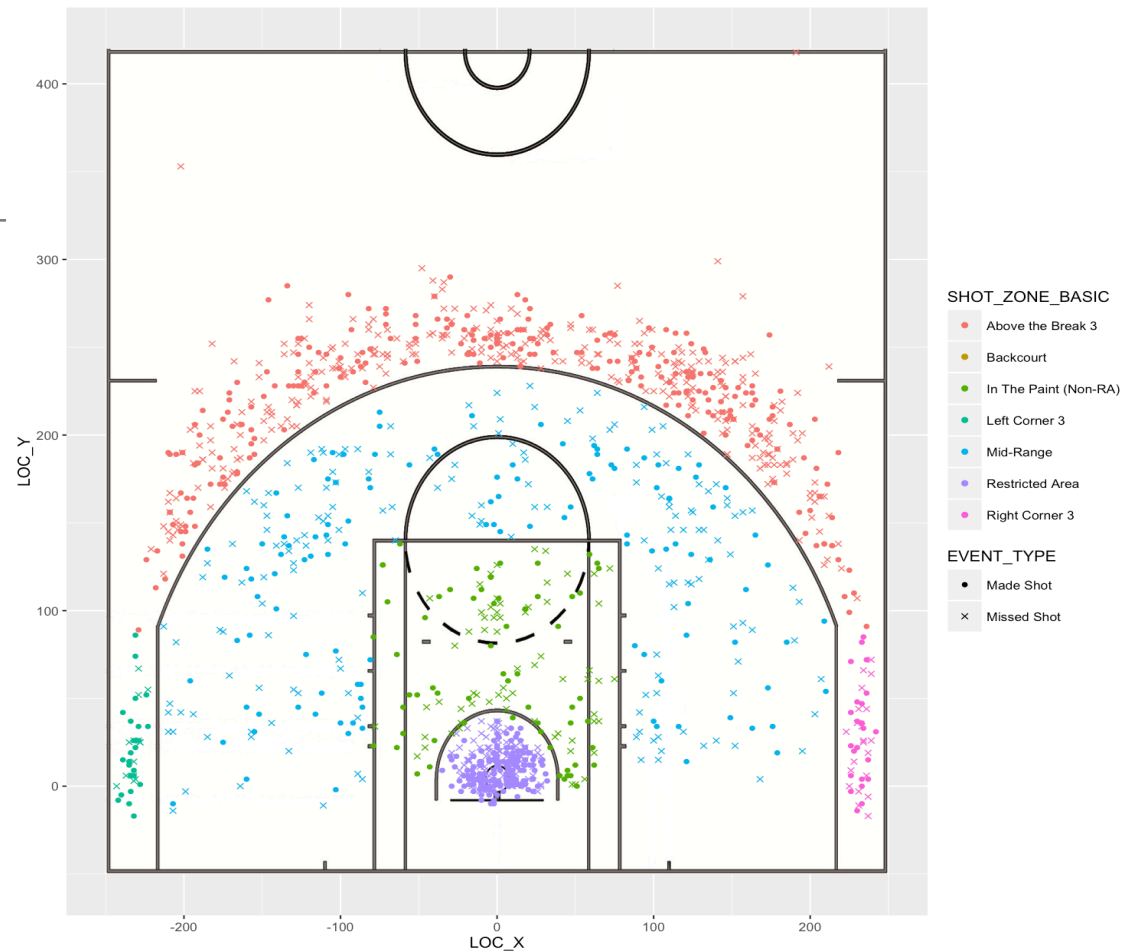
Anna Matlin
amatlin@princeton.edu

End goal

Shown on right: visualized shot data for Steph Curry of the Golden State Warriors

By the end of this workshop, we will have this plot displayed in a Shiny application, with player options and information

Resource dump

➢ All code used in this workshop: https://github.com/amatlin/NBAvis

➢ Stats data retrieved from http://www.nba.com/warriors/stats

➢ Scraping:
  ➢ https://www.crummy.com/software/BeautifulSoup/bs4/doc/
  ➢ http://web.stanford.edu/~zlotnick/TextAsData/Web_Scraping_with_Beautiful_Soup.html   (tutorial)

➢ Shiny:
  ➢ http://shiny.rstudio.com/
  ➢ http://shiny.rstudio.com/articles/cheatsheet.html

➢ Download python: https://www.python.org/downloads/

➢ Download R and Rstudio:
  ➢ http://lib.stat.cmu.edu/R/CRAN/
  ➢ https://www.rstudio.com/products/rstudio/download/

➢ ggplot2:
  ➢ http://ggplot2.org/
  ➢ https://www.rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf

➢ **Lots of code used** from http://thedatagame.com.au/2015/09/27/how-to-create-nba-shot-charts-in-r/

Step 1: Have Python and R installed

Python needed for web scraping

R, RStudio needed for visualization + Shiny

Step 2: go to Github repository

https://github.com/amatlin/NBAvis

→ open README and follow instructions for installation of necessary packages

Step 3: Scraping

Go to http://www.nba.com/warriors/stats to get a look for what we will be scraping.

Player IDs come from image links in HTML.

Find image of player, right click and select "Inspect"

Frameworks/APIs we will be using to scrape:

→ Beautiful Soup

```
<td class="stl">0.4</td>
<td class="tov">0.9</td>
<td class="pf">2.3</td>
</tr>
▼<tr no_striping="1">
  ▼<td class="player_name">
    ▼<div class="player-name__inner-wrapper">
        <img src="http://stats.nba.com/media/players/
        230x185/203110.png" width="60px" alt>
      ▶<span class="playerInfo">…</span>
    </div>
  </td>
  <td class="gp">74</td>
  <td class="pts">13.8</td>
  <td class="fgm">4.9</td>
  <td class="fg_pct">48.6%</td>
  <td class="fg3_pct">39%</td>
```

What is a Shiny app?

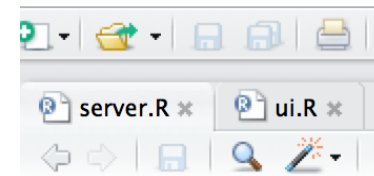Check out http://shiny.rstudio.com/ for documentation + tutorials

→ "**A web application framework for R;** Turn your analyses into interactive web applications; No HTML, CSS, or JavaScript knowledge required"

→ each app is made containing two files
- **server.R**: behind the scenes calculations
- **ui.R**: what the user sees

→ designed for development in RStudio

→ Terms to keep in mind: *render, reactive, input/output*

Step 4: Start making our app

Step 5: ggplot2

For more info check out

➢ http://ggplot2.org/

➢ https://www.rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf

"ggplot2 is a plotting system for R, based on the grammar of graphics, **which tries to take the good parts of base and lattice graphics and none of the bad parts**. It takes care of many of the fiddly details that make plotting a hassle (like drawing legends) as well as providing a powerful model of graphics that makes it **easy to produce complex multi-layered graphics**."