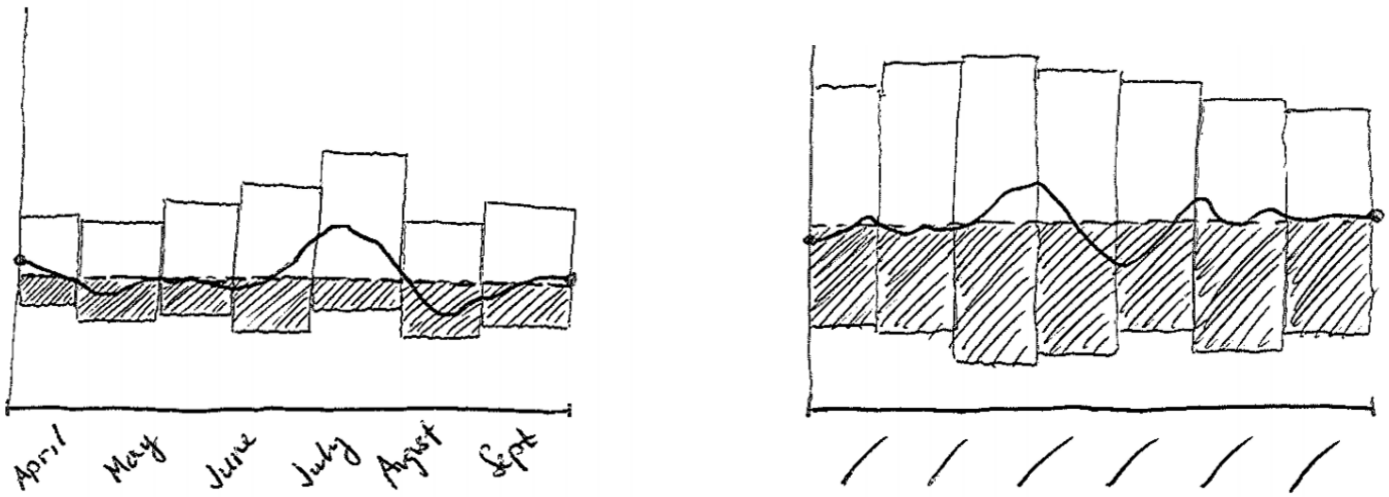
Idea 1: Exploring Sports Fan Loyalty

- **Problem:**
  - Which cities/teams/sports have the most loyal fans?
    - Which cities have the most die-hard fans vs fair-weather fans? By sport? In general? Look at team records.
    - Which teams' fans are most polarized (high positive/high negative)?
    - Teams with wide vs narrow fan base (geotags)?
    - Most hated teams (good record? bad record?)
- **Data:**
  - Twitter API for sentiment analysis associated with certain hashtags (#redsox, #bruins, etc.).
  - Twitter geotags
  - Team yearly performance records (scrape or find data set)
- **Hypothesis:**
  - Some effect of sport, city, and team on robustness of fan base.

Idea 1: Exploring Sports Fan Loyalty



Idea 2: Predicting MLB Home Run Derby Winner

- **Problem:**
  - Who will win the MLB Home Run Derby?
- **Data:**
  - Yearly regular-season player stats (scrape? data set?)
  - Previous derby winners by year, pitcher, HRs each round, etc.: (sports.newsday.com/long-island/data/baseball/all-star-game/home-run-derby/)
- **Hypothesis:**
  - Use ML to find effect of certain regular-season stats on likelihood of winning Home Run Derby. Likely significant factors: batting average, avg home runs per season, etc.

Idea 3: Map-based exploration of who watches which TV shows?

- **Problem:**
  - What does your geographic/demographic info say about the TV shows you watch (read: Tweet about)?
- **Data:**
  - Twitter API for geotagging TV-related hashtags.
  - Overlay US Census data by city our county. E.g., unemployment status by county: http://www.census.gov/people/laborforce/about/acs_employ.html.
  - Other possible overlays: race (http://demographics.coopercenter.org/DotMap/), average income, crime rate, etc.
- **Hypothesis:**
  - Expect to find some effect of demographic information on TV shows you Tweet about, through the lens of location.