

MIS 0855 Spring 2015 – Data Science

Day 5 – Identifying Data Sources

Min-Seok Pang

**Management Information Systems
Fox School of Business, Temple University
minspang@temple.edu**

Jan. 23rd, 2015

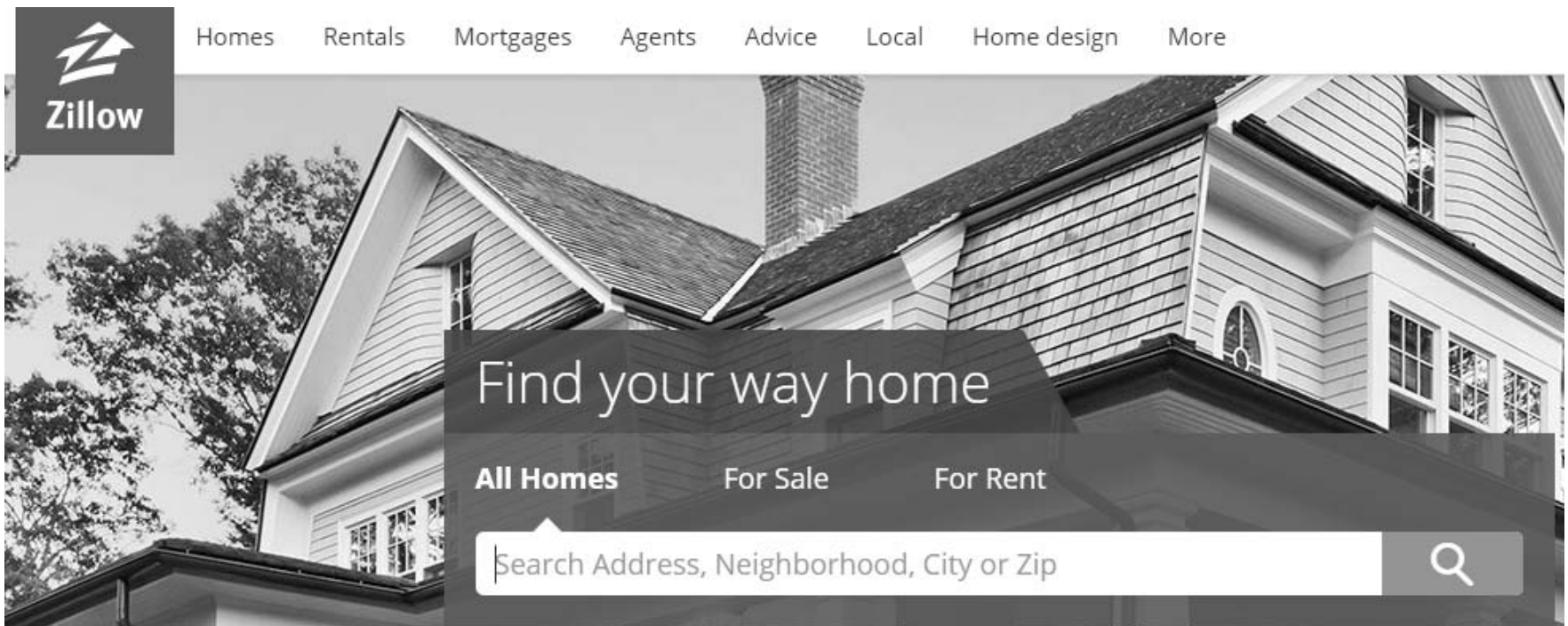
How to find data? (1/2)

- You have a decision to make.
- You have theory in mind and hypotheses to test.
- Where would you find data for it?



How to find data? (2/2)

- “Houses in Philly Neighborhood A are bigger than in Neighborhood B.”
- Where can you find data to test this hypothesis?



Zillow API



```

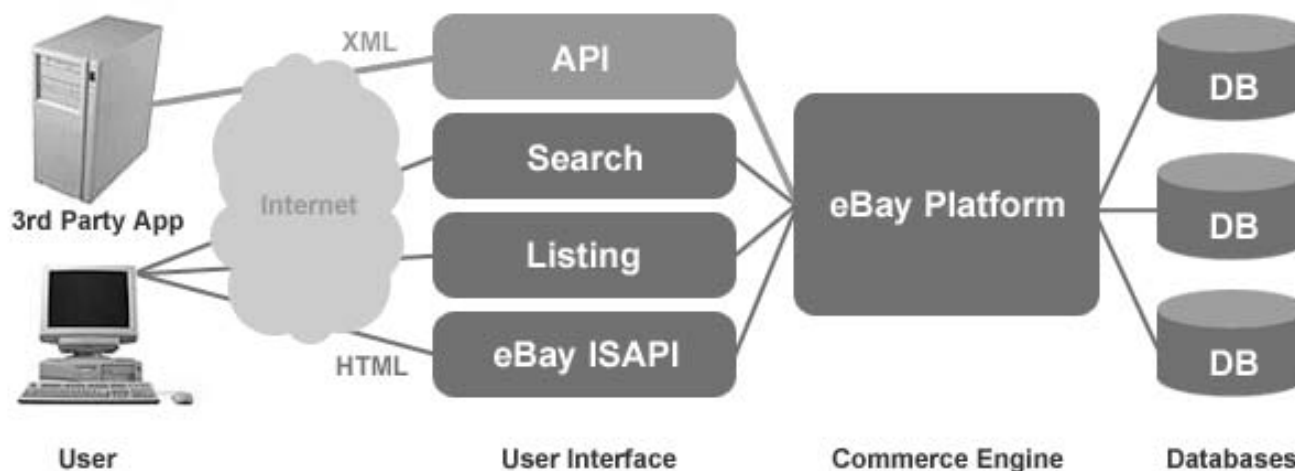
▼ <SearchResults:searchresults xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.zillow.com/static/xsd/SearchResults.xsd http://www
  ▼ <request>
    <address>500 Walnut St</address>
    <citystatezip>Philadelphia PA</citystatezip>
  </request>
  ▼ <message>
    <text>Request successfully processed</text>
    <code>0</code>
  </message>
  ▼ <response>
    ▼ <results>
      ▼ <result>
        <zpid>2110350697</zpid>
        ▼ <links>
          ▼ <homedetails>
            http://www.zillow.com/homedetails/400-Walnut-St-APT-3F-Philadelphia
          </homedetails>
          <mapthishome>http://www.zillow.com/homes/2110350697_zpid/</mapthishome>
          <comparables>http://www.zillow.com/homes/comps/2110350697_zpid/</comparables>
        </links>
        ▼ <address>
          <street>400 Walnut St APT 3F</street>
          <zipcode>19106</zipcode>
          <city>Philadelphia</city>
          <state>PA</state>
          <latitude>39.947258</latitude>
          <longitude>-75.148101</longitude>
        </address>
        ▼ <zestimate>
          <amount currency="USD">658805</amount>
          <last-updated>12/21/2014</last-updated>
          <oneWeekChange deprecated="true"/>
          <valueChange duration="30" currency="USD">-142820</valueChange>
          ▼ <valuationRange>
  
```



0 days on Zillow

API – Application Programming Interface

- Software that exposes internal data to the outside world.
- People write programs that “talk” to the API so it will fetch data for them.
- Keeps people from accessing the data directly, keeping it safe.





Find tacos, cheap dinner, Max's

Near New York, NY



Sign Up

Home About Me Write a Review Find Friends Messages Talk Events

Log In

The Joyce Theater

★★★★☆ 38 reviews [Details](#)

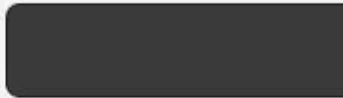
★ Write a Review

📷 Add Photo

➦ Share

🔖 Bookmark

Performing Arts [Edit](#)

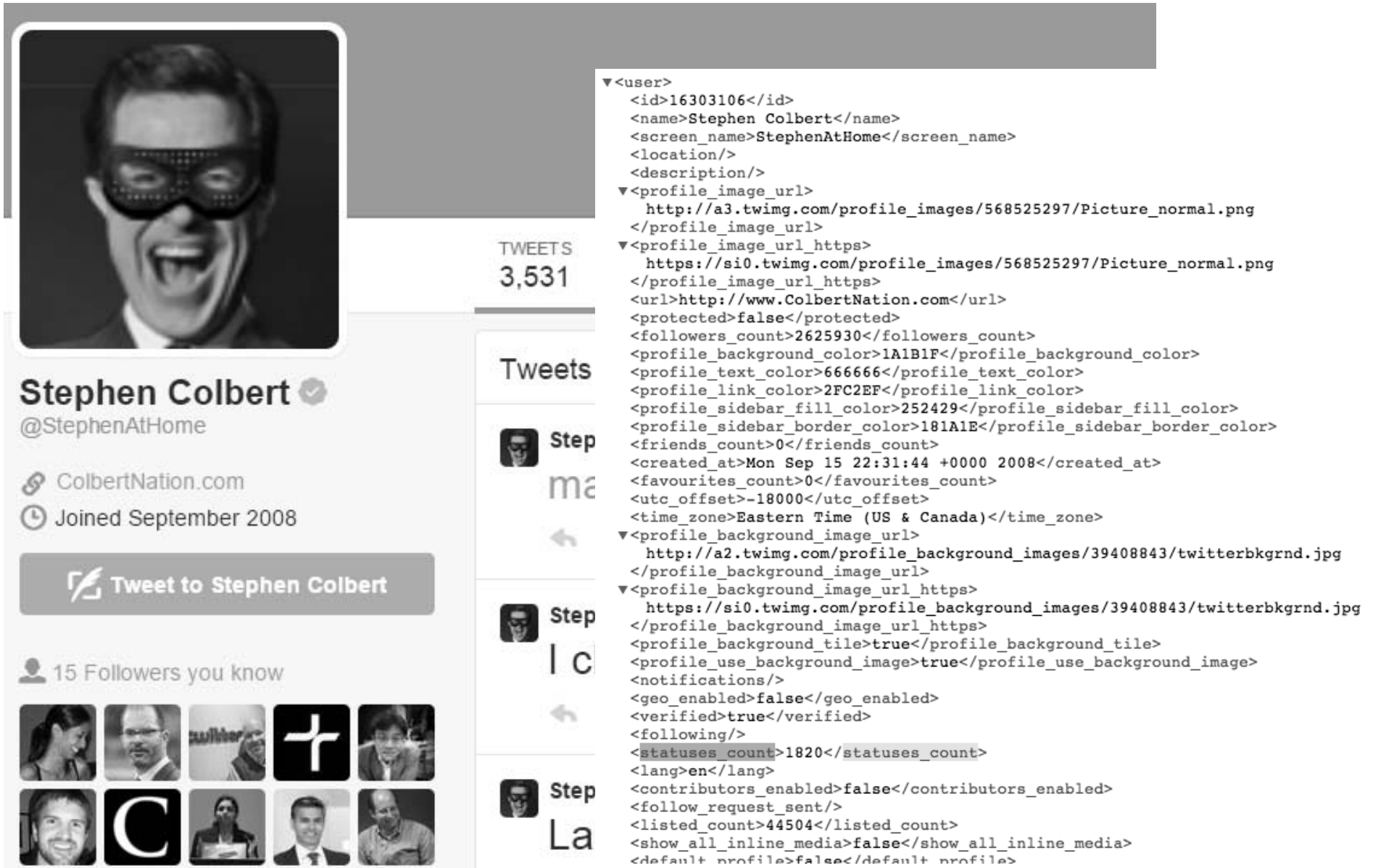


175 8th Ave
New York, NY 10011
b/t 19th St & 18th St in Ch
[Get Directions](#)
🚗 **A C E** 14 St. and 2
📞 (212) 691-9740
🌐 joyce.org

```
=> {"region"=>
  {"span"=>
    {"latitude_delta"=>0.2769805571191881,
      "longitude_delta"=>0.0645380999999294},
    "center"=>{"latitude"=>40.703339046764, "longitude"=>-73.9729782}},
  "total"=>17,
  "businesses"=>
    [{"is_claimed"=>true,
      "rating"=>4.5,
      "mobile_url"=>"http://m.yelp.com/biz/the-joyce-theater-new-york-2",
      "rating_img_url"=>
        "http://s3-media2.ak.yelpcdn.com/assets/2/www/img/99493c12711e/ico/stars/v1/stars_4_half.png",
      "review_count"=>34,
      "name"=>"The Joyce Theater",
      "snippet_image_url"=>
        "http://s3-media1.ak.yelpcdn.com/photo/DBRJa6-PpFWCQqP95jqq/ms.jpg",
      "rating_img_url_small"=>
        "http://s3-media2.ak.yelpcdn.com/assets/2/www/img/a5221e66bc70/ico/stars/v1/stars_small_4_half.png",
      "url"=>"http://www.yelp.com/biz/the-joyce-theater-new-york-2",
      "phone"=>"2126919740",
      "snippet_text"=>
        "The Joyce comes through again - this time we went to see MOMIX the day after Christmas and it was a spectacular show. But here's where the Joyce shines: the...",
      "image_url"=>
        "http://s3-media4.ak.yelpcdn.com/bphoto/0PlmKtRE85LGGp9lx3uIdw/ms.jpg",
      "categories"=>[["Performing Arts", "theater"]],
      "display_phone"=>"+1-212-691-9740",
      "rating_img_url_large"=>
        "http://s3-media4.ak.yelpcdn.com/assets/2/www/img/9f83790ff7f6/ico/stars/v1/stars_large_4_half.png",
      "id"=>"the-joyce-theater-new-york-2",
      "is_closed"=>false,
      "location"=>
        {"cross_streets"=>"19th St & 18th St",
          "city"=>"New York",
          "display_address"=>
            ["175 8th Ave",
              "(b/t 19th St & 18th St)",
              "Chelsea",
              "New York, NY 10011"],
          "neighborhoods"=>["Chelsea"],
          "postal_code"=>"10011",
          "country_code"=>"US",
          "address"=>["175 8th Ave"],
          "state_code"=>"NY"}},
      {"is_claimed"=>true,
        "rating"=>4.0,
        "mobile_url"=>"http://m.yelp.com/biz/symphony-space-new-york",
        "rating_img_url"=>
          "http://s3-media4.ak.yelpcdn.com/assets/2/www/img/c2f3dd9799a5/ico/stars/v1/stars_4.png",
        "review_count"=>36,
        "name"=>"Symphony Space",
        "snippet_image_url"=>
          "http://s3-media1.ak.yelpcdn.com/photo/emmcnQilGJrcw895TVaitw/ms.jpg",
        "rating_img_url_small"=>
```

<http://www.yelp.com/biz/the-joyce-theater-new-york-2>

<https://medium.com/@jmhinc/using-oauth-for-yelp-api-requests-1e286c83e72e>



The image shows a screenshot of Stephen Colbert's Twitter profile. On the left is his profile picture (a black and white photo of him wearing glasses and shouting). Below it is his name "Stephen Colbert" with a verified badge, his handle "@StephenAtHome", and links to "ColbertNation.com" and "Joined September 2008". A button says "Tweet to Stephen Colbert". Below that, it says "15 Followers you know" and shows a grid of follower avatars. On the right, the "Tweets" section shows a list of tweets, with the top one partially visible. Overlaid on the right side of the screenshot is a JSON data dump for the user's profile information.

```

<user>
  <id>16303106</id>
  <name>Stephen Colbert</name>
  <screen_name>StephenAtHome</screen_name>
  <location/>
  <description/>
  <profile_image_url>
    http://a3.twimg.com/profile_images/568525297/Picture_normal.png
  </profile_image_url>
  <profile_image_url_https>
    https://si0.twimg.com/profile_images/568525297/Picture_normal.png
  </profile_image_url_https>
  <url>http://www.ColbertNation.com</url>
  <protected>>false</protected>
  <followers_count>2625930</followers_count>
  <profile_background_color>1A1B1F</profile_background_color>
  <profile_text_color>666666</profile_text_color>
  <profile_link_color>2FC2EF</profile_link_color>
  <profile_sidebar_fill_color>252429</profile_sidebar_fill_color>
  <profile_sidebar_border_color>181A1E</profile_sidebar_border_color>
  <friends_count>0</friends_count>
  <created_at>Mon Sep 15 22:31:44 +0000 2008</created_at>
  <favourites_count>0</favourites_count>
  <utc_offset>-18000</utc_offset>
  <time_zone>Eastern Time (US & Canada)</time_zone>
  <profile_background_image_url>
    http://a2.twimg.com/profile_background_images/39408843/twitterbkgrnd.jpg
  </profile_background_image_url>
  <profile_background_image_url_https>
    https://si0.twimg.com/profile_background_images/39408843/twitterbkgrnd.jpg
  </profile_background_image_url_https>
  <profile_background_tile>true</profile_background_tile>
  <profile_use_background_image>true</profile_use_background_image>
  <notifications/>
  <geo_enabled>>false</geo_enabled>
  <verified>true</verified>
  <following/>
  <statuses_count>1820</statuses_count>
  <lang>en</lang>
  <contributors_enabled>>false</contributors_enabled>
  <follow_request_sent/>
  <listed_count>44504</listed_count>
  <show_all_inline_media>>false</show_all_inline_media>
  <default_profile>>false</default_profile>
  
```

<https://twitter.com/StephenAtHome>
<http://ruby.bastardsbook.com/chapters/conclusion/>

Open Government Data

- A government is a collector, generator, and consumer of data.
 - What kind?
- Data from any government service and function is public assets.
- The Freedom of Information Act (FOIA)



<http://gsablogs.gsa.gov/innovation/2013/08/27/open-data-the-next-phase-in-the-technology-revolution/>

