**DATA SKILLS/SECTION 2**

**Class Notes for Wednesday, October 14**

**Robert Gebeloff**

**rmg2196@columbia.edu**

**AGENDA:**

1. **Your final “exam”**
2. **Exercise 1 – Getting HTML into XLS with Web queries**
3. **Exercise 2 – Wrangling PDF files**
4. **Exercise 3 – Scraping With the Chrome Scraper**
5. **Exercise 4 – An XML Miracle**

**Scheduling Note:**

**REMINDER: NO CLASS NEXT WEEK, we resume October 14**

**So:**

**Class 4 – October 14**

**Class 5 – October 21**

**Class File Repository:**

[**https://github.com/sarahcnyt/data-journalism**](https://github.com/sarahcnyt/data-journalism)