

INTRO to DATA SCIENCE

LECTURE 3: PYTHON, PANDAS, MATPLOTLIB

- 1. INTRO, UNIX**
- 2. SQL, PYTHON**
- 3. PYTHON, PANDAS, MATPLOTLIB**
- 4. MORE GATHERING (WEB SCRAPING, APIS) AND
HOW TO PRESENT YOUR INSIGHTS**

CLOSED BY ASSIGNMENT #1

I. REVIEW LAST CLASS

- GENERAL QUESTIONS ABOUT SQL & PYTHON**
- EXERCISES DATA EXPLORATION IN PYTHON**

II. HEADS-UP ASSIGNMENT #1

II. PANDAS

III. VISUALIZATONS

INTRO TO DATA SCIENCE

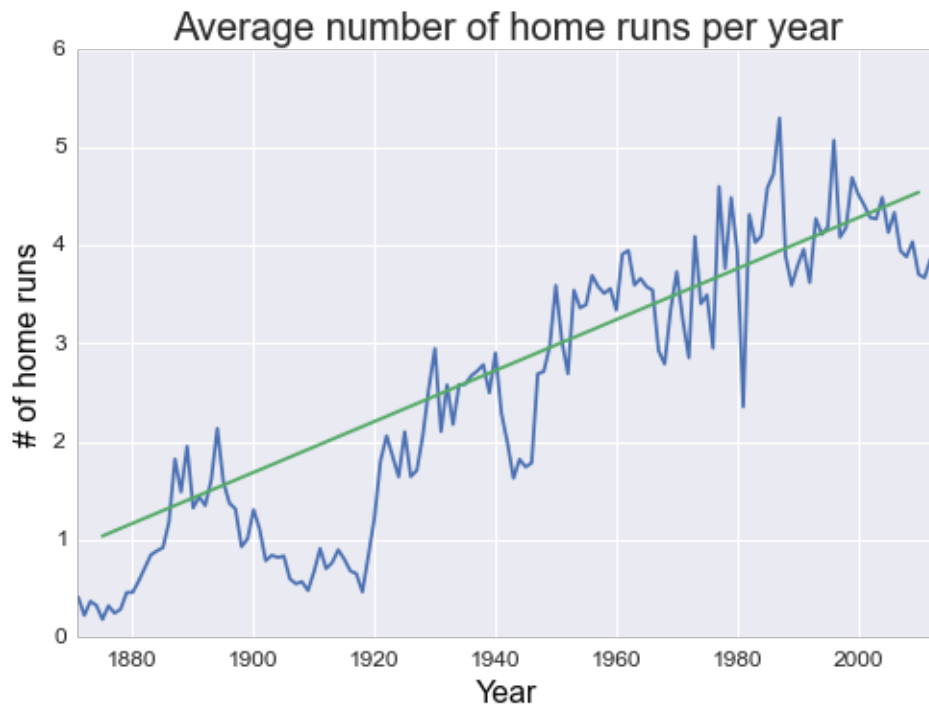
II. ASSIGNMENT #1

- GATHER A DATASET
- POSE A FEW QUESTIONS
- PROCESS THE DATA IN PYTHON AND PANDAS
- GENERATE DESCRIPTIVE STATISTICS
- CREATE VISUALIZATIONS
- SUBMIT YOUR WORK (DUE **SUN 7/17 MIDNIGHT**)
- PRESENT IN CLASS (**TUE 7/14**)

III. VISUALIZATIONS

HOME RUNS GET MORE AND MORE COMMON

7



- › One message per slide
- › Conclusion in header (so what?)
- › Clear labels in chart (title, axes)
- › Trend line emphasizes message
- › Source at the bottom

Think of your narrative:
your slides should form a story

- Go to github.com/ga-students/DAT-23-NYC
- Scroll down to lesson #3