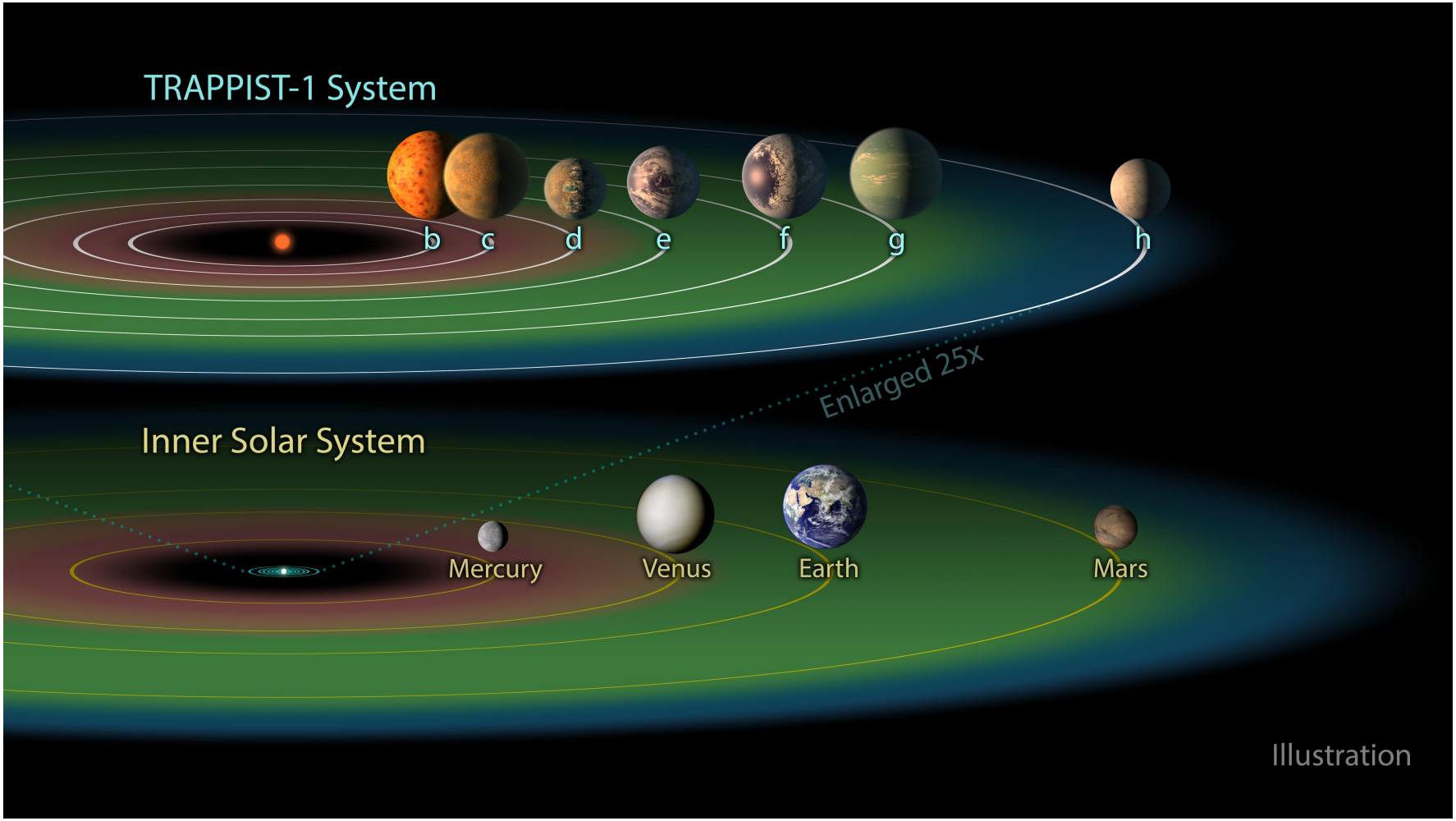
YData: ExoStatistics

Exploring Extrasolar Planets with Data Science

S&DS 170/570, ASTR 445/545

Lecture 01
Introduction

Jessi Cisewski-Kehe
Statistics & Data Science, Yale University
Spring 2019



TRAPPIST-1 (star): radius $\approx$ 0.121 Solar radius (slightly larger than Jupiter's radius of about 0.10 Solar radius)
mass $\approx$ 0.089 Solar mass (about 93 times more massive than Jupiter)
Mission: Spitzer Space Telescope (NASA / JPL / Caltech)
Instrument: TRAPPIST
Image credit: https://www.jpl.nasa.gov/spaceimages/details.php?id=pia21424

Staff

Instructor
Jessi Cisewski-Kehe

TF
Xin Xu

Course website

https://canvas.yale.edu

This term we will be using Piazza for class discussion. The system is highly catered to getting you help fast and efficiently from classmates, the TA, and myself. Rather than emailing questions to the teaching staff, I encourage you to post your questions on Piazza.

Find our class page at: https://piazza.com/yale/spring2019/sds170/home

YData website: http://ydata123.org

YData Seminar Courses

- Data science is driven by applications
- Every data-driven subject brings new challenges
- YData Seminars are small, independent courses taught by Yale faculty who are excited to share their expertise

Course Structure

- One 110-minute meeting per week (Tuesdays 3:30 - 5:20PM)

  $\longrightarrow$ Lecture and Lab time
- Weekly lab assignments (started, and possibly completed, during class)
- Two midterms during class: **2/19 and 4/2**
- Office hours are posted on our class Canvas site.
- One final project (no final written exam)

Details can be found on the course Canvas site.

Course Textbook

Computational and Inferential Thinking:
The Foundations of Data Science

By Ani Adhikari and John DeNero (Adhikari and DeNero, 2018)

Freely available at https://www.inferentialthinking.com

Grade distribution

Undergraduate level		Graduate level	
Lab assignments (weekly)	40%	Lab assignments (weekly)	40%
Midterm exam 1	15%	Midterm exam 1	15%
Midterm exam 2	15%	Midterm exam 2	15%
Final project	20%	Final project and presentation	20%
Participation	10%	Participation	10%
Total	100%	Total	100%

Grades

Course grades will appear on the course website. Each student is responsible for verifying his or her recorded scores during the semester.

Your overall course score will be determined as a weighted average of each element as noted on the previous slide. A letter grade will be assigned based on:

A: 93-100	**A-:** 90-93	**B+:** 87-90	**B:** 83-87	**B-:** 80-83
C+: 77-80	**C:** 73-77	**C-:** 70-73	**D:** 60-70	**F:** Below 60

Honors: 90 - 100	**High Pass:** 80 - 90	**Pass :** 70 - 80	**Fail:** Below 70

Although we strive for consistency and accuracy in grading, we understand that grading errors can occur.

- We will gladly correct all errors in tabulation or overlooked material.
- All regrading requests must be accompanied by a written statement carefully highlighting and explaining the items that were mis-graded. *Note that regrading requests can end in a positive, negative, or no change in points.*
- All regrade requests should be submitted to the instructor within one week of when the graded work is returned.

Participation

- Evaluated based on Attendance, listening/asking/answering questions

  Actively working on the lab assignments during the lab portion of the class

  Usage of Piazza
- Half of the total participation portion of the grade will be evenly distributed across the class meetings minus the one allowed absence (e.g. due to illness)

  The other half will be based on Piazza participation
- Students who (i) attend class weekly and participate in the class activities, and (ii) post on Piazza about once every two weeks will receive full credit.

Circumstances can arise during a semester when attendance is not possible (e.g. due to health issues) and so a student can miss one class without it affecting the participation portion of the grade; students who miss a class are still expected to complete the lab assignment and turn it in on time.

Final project

- The final project is an opportunity for you to explore, in more detail, a question of interest related to exoplanets and data science. It is expected that the topic of the question will be related to exoplanets and the project will include an analysis using data science methods.
- The final project will culminate in a 5 - 10 page written report introducing your question, describing the methodology you employ to answer the question, a discussion of the results, and finally the conclusions you can draw from the analysis.
- In addition to the written report, graduate students will present the results of the final project during one of the final class periods (the length will depend on the number of graduate students enrolled).

We will discuss this in more detail during the semester!

Honor Code

- You are encouraged to be helpful to your classmates and to work together, but the work you turn in must be your own. Any student who turns in work for credit that is identical, or similar beyond coincidence, to that of another student may face appropriate disciplinary action at the department, college, or university level. Cheating and/or plagiarism will not be tolerated.

- If you get ideas or words from a website, journal article, book, another person, etc., cite the source in your work at the location where you use the idea. Then include a bibliography or list of sources cited at the end of your document.

- For more information about plagiarism and how to avoid it, please see Yale's Center for Teaching and Learning website on the topic:

  https://poorvucenter.yale.edu/writing/using-sources/understanding-and-avoiding-plagiarism

Course schedule

On our Canvas site, you will find the following calendar with links to relevant documents

Course Schedule

Date	Topic	Reading	Lab assignment
Week 1 (1/15)	Introduction to Exoplanets Getting started with Python		Lab 00 (not graded)
Week 2 (1/22)	Detecting Exoplanets: Transit Method		Lab 01 (due 1/28)
Week 3 (1/29)	Exploring Exoplanet Populations		Lab 02 (due 2/4)
Week 4 (2/5)	Detecting exoplanets: Radial Velocity Method		Lab 03 (due 2/11)
Week 5 (2/12)	Stellar Activity		Lab 04 (due 2/19)
Week 6 (2/19)	Midterm Exam 1		Lab 05 (due 3/4)

General comments

- This is a half-credit course so the overall workload will reflect this
- General goal is to have extra practice using skills learned in YData, but for a particular field (exoplanets!)
- We will learn about specific data science techniques for detecting and characterizing exoplanets; these topics will not generally be covered in the main YData course
- I am a statistician/data scientist, not a trained astronomer

What is an exoplanet?

A planet orbiting a star outside our Solar System.

What is an exoplanet?

A planet orbiting a star outside our Solar System.

So...what's a planet?

Ancient Greek planētēs = "wanderer"

In 2006, the International Astronomical Union (IAU) decided on the following definition:

Planet

A celestial body that

1. is in orbit around the Sun
2. has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape, and
3. has cleared the neighbourhood around its orbit.

http://www.iau.org/public_press/news/detail/iau0603/

Hence the Solar System has 8 planets:

Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus and Neptune

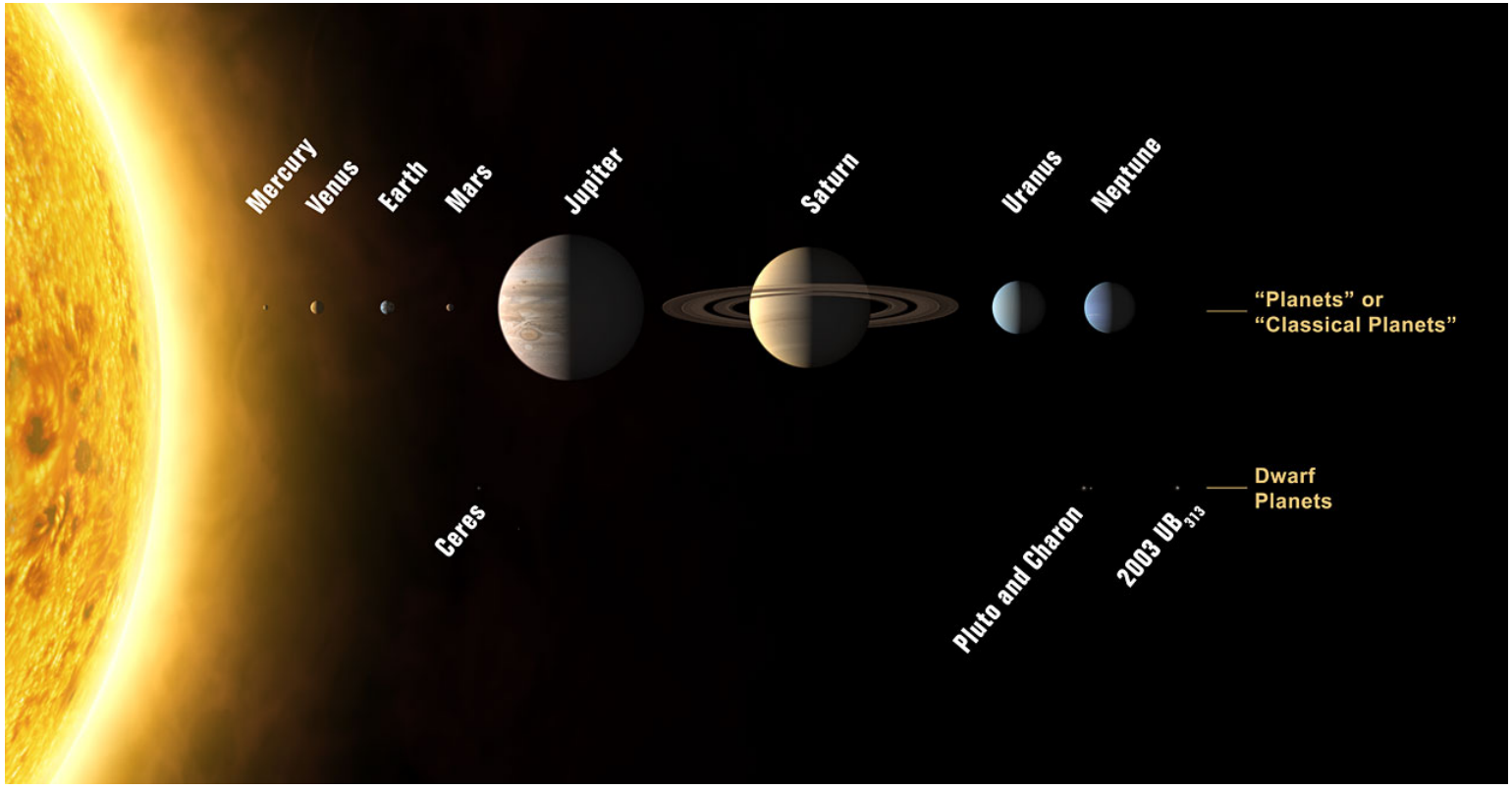
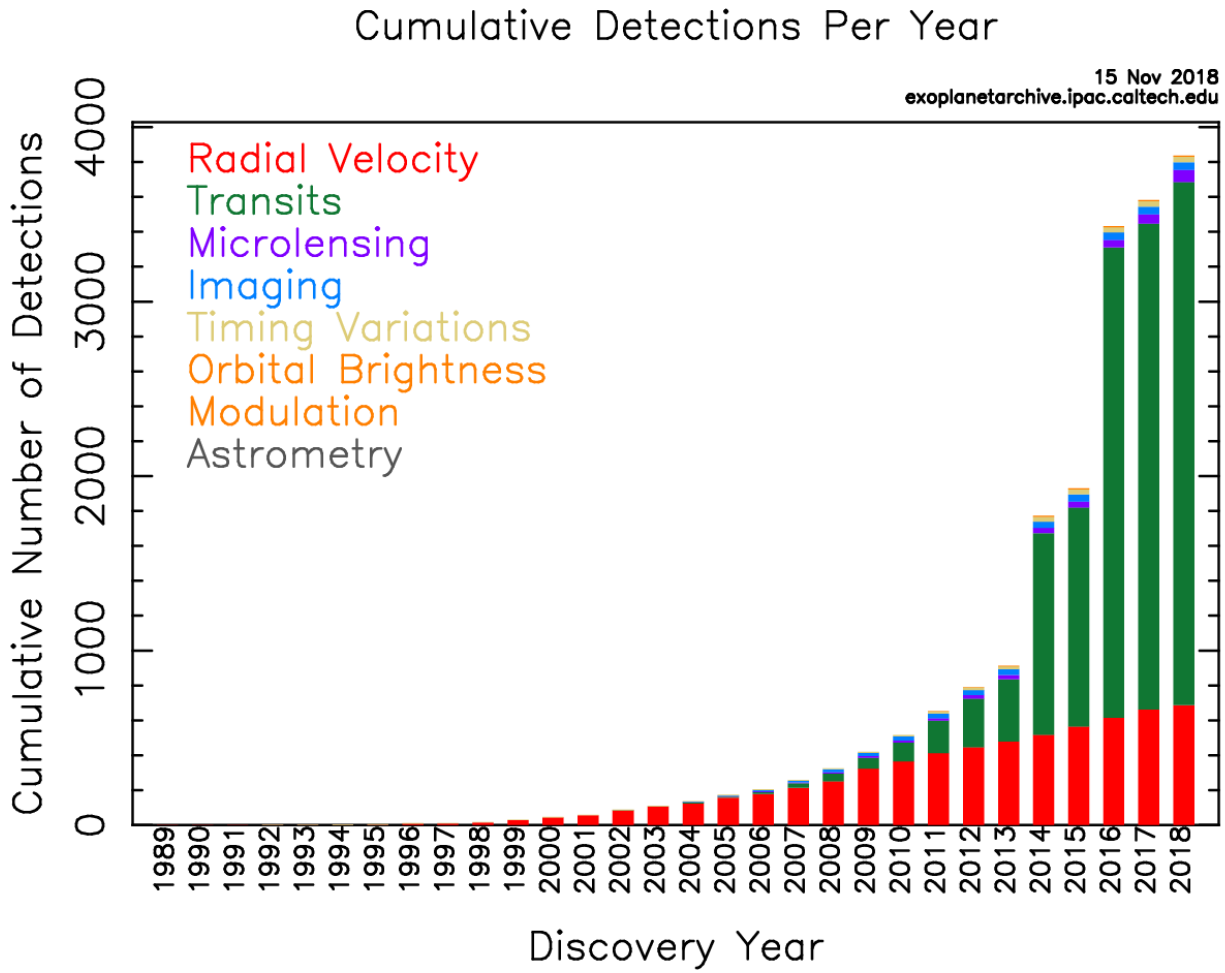


Figure credit: http://www.iau.org/public_press/news/detail/iau0603/



http://exoplanets.org and Debra Fischer

Main course topics

- **Detection of exoplanets**

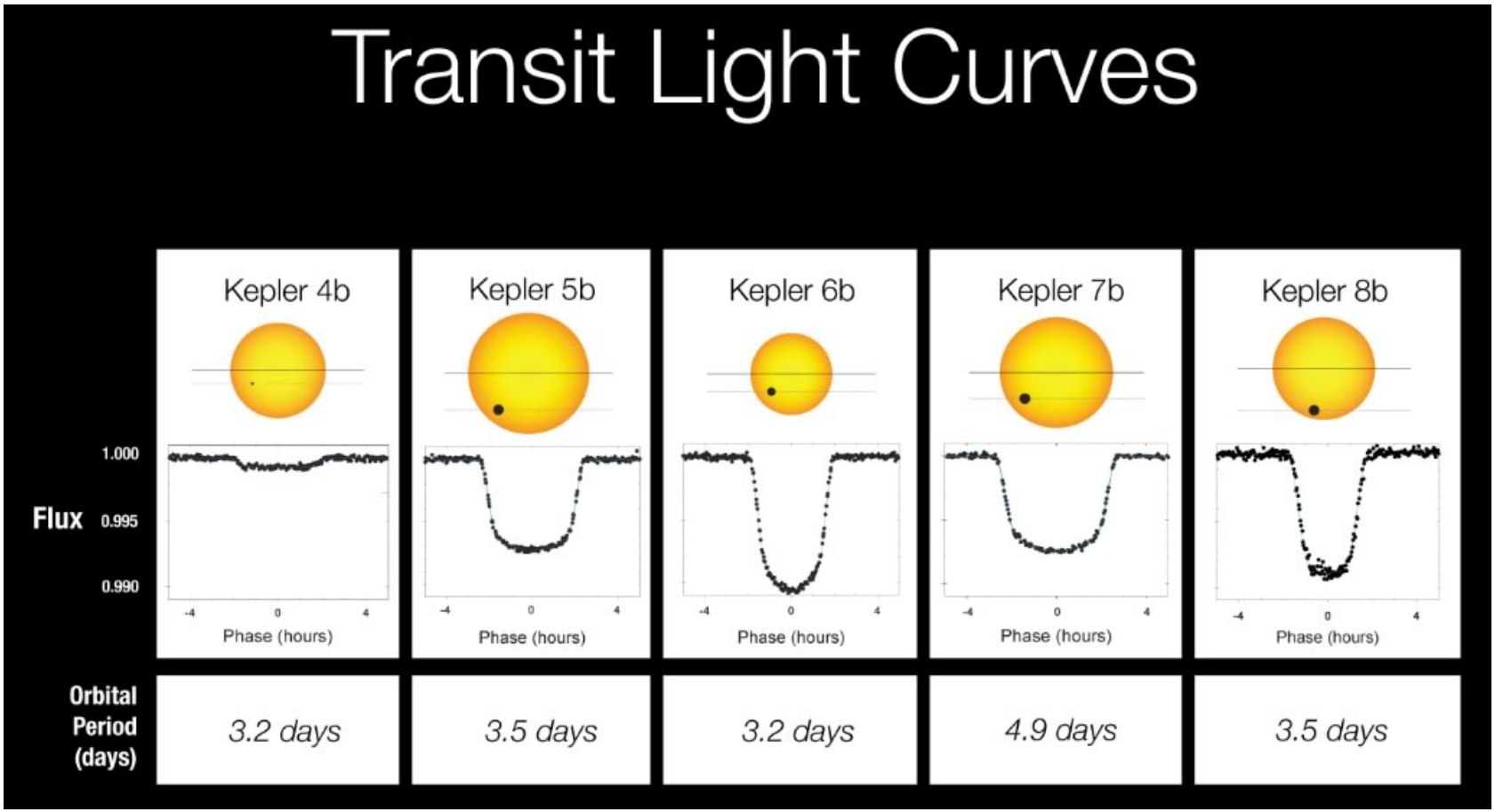
  $\longrightarrow$ We will learn about the popular methods of detection and we will implement them on real data

- **Characterization of exoplanets**

  $\longrightarrow$ Once we find exoplanets, we want to learn about them...what are they like? what is the distribution of their masses? how many of each type of exoplanet have we found? how many orbit a star?

Transit method

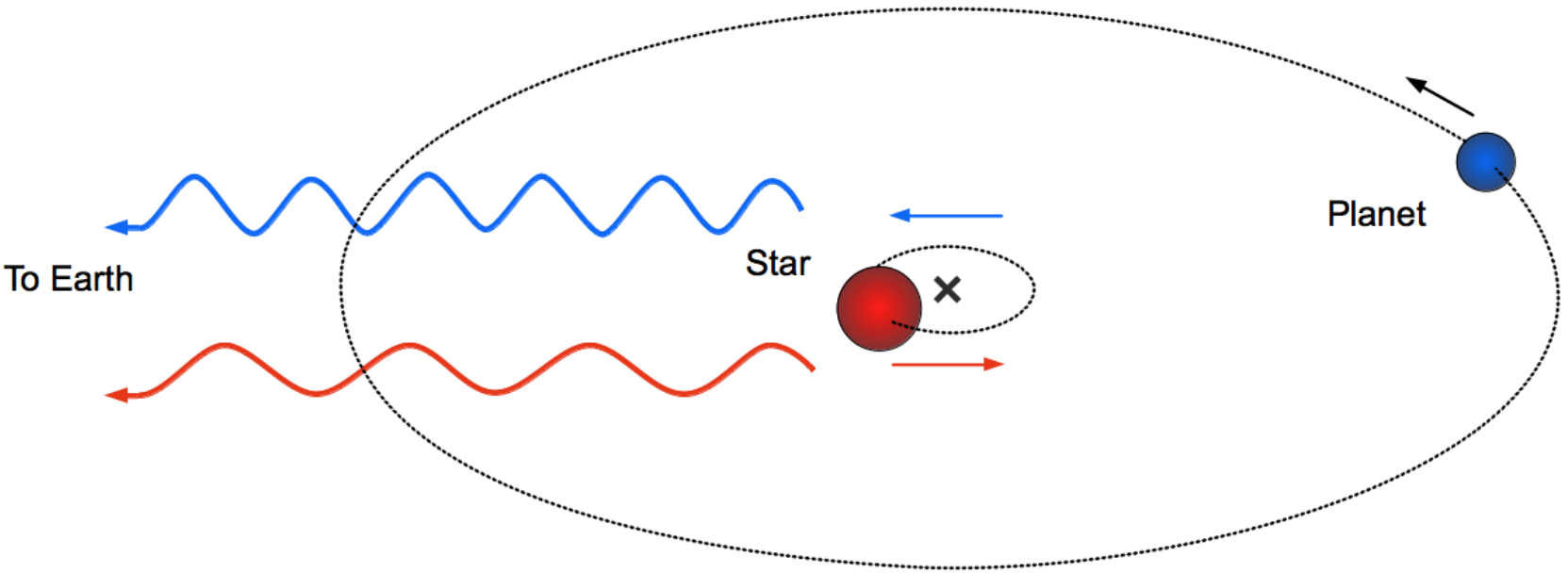
- Exoplanet(s) orbiting a host star can lead to dips in light curves when exoplanets transit



Plot: NASA/Kepler Mission

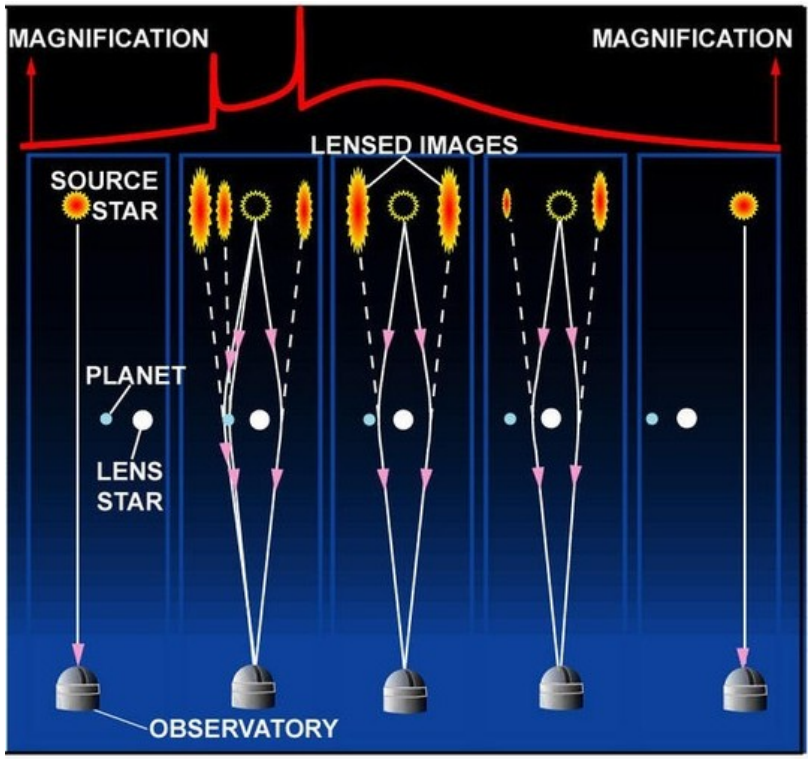
Radial Velocity (RV) method

- Exoplanet(s) orbiting a host star can lead to **Doppler-shifted** stellar spectra



Plot: https://upload.wikimedia.org/wikipedia/commons/

Gravitational Microlensing



Plot: OGLE

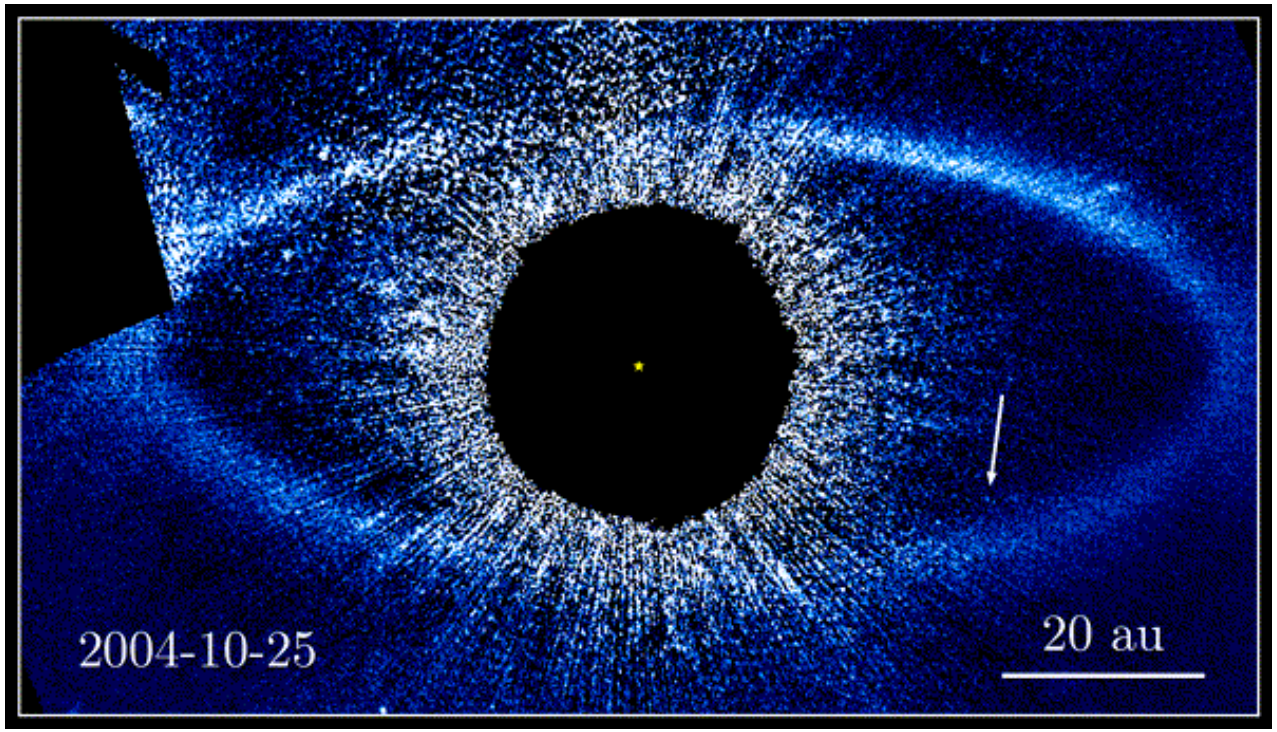
The microlensing process in stages, from right to left.

The lensing star (white) moves in front of the source star (yellow) magnifying its image and creating a microlensing event.

In the fourth image from the right the planet adds its own microlensing effect, creating the two characteristic spikes in the light curve.

Description from http://www.planetary.org/explore/space-topics/exoplanets/microlensing.html

Imaging



- Fomalhaut b (aka "Dagon")
- Discovered November 13, 2008 using the Hubble Space Telescope (Kalas et al., 2008)
- Orbital period of about 1700 years.

Credit: Jason Wang/Paul Kalas; UC Berkeley;

https://sciencesprings.wordpress.com/2018/05/02/from-many-worlds-exoplanet-fomalhaut-b-on-the-move/

Exoplanet naming conventions

- International Astronomical Union statement: https://www.iau.org/public/themes/naming_exoplanets/

- Scientific nomenclature

  1) A proper noun or abbreviation

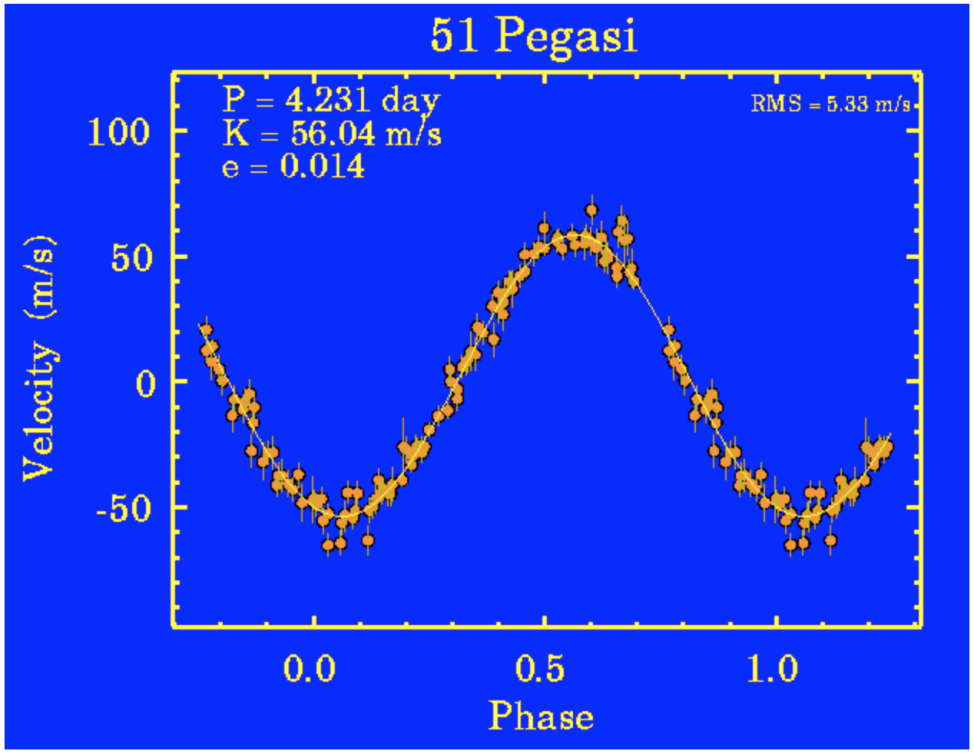
  → A common source is an exoplanet's host star's widely recognized, common or astronomical catalogue name. Alternatively, the scientific instrument/project that made discovery

  2) Followed by a lowercase letter

  → Indicates the order of the planet's discovery around its host star (first discovered is designated b; the second, c; the third d; etc.).

  E.g. 51 Pegasi b, Kepler-90 b, c, d, e, f, g, h, and i

51 Pegasi b



exoplanets.org

- 51 Pegasi b, discovered using RV Method (Mayor and Queloz, 1995)
- First exoplanet discovered orbiting a main-sequence star (fuse hydrogen into helium in core, the Sun is a main-sequence star)
- Minimum mass of 150 Earth masses (about half the mass of Jupiter)
- Semimajor axis ≈ 0.05 AU (closer to host than Mercury is to the Sun!)
- Orbital period of 4.23 days, and an eccentricity of about 0 (i.e. an almost circular orbit)

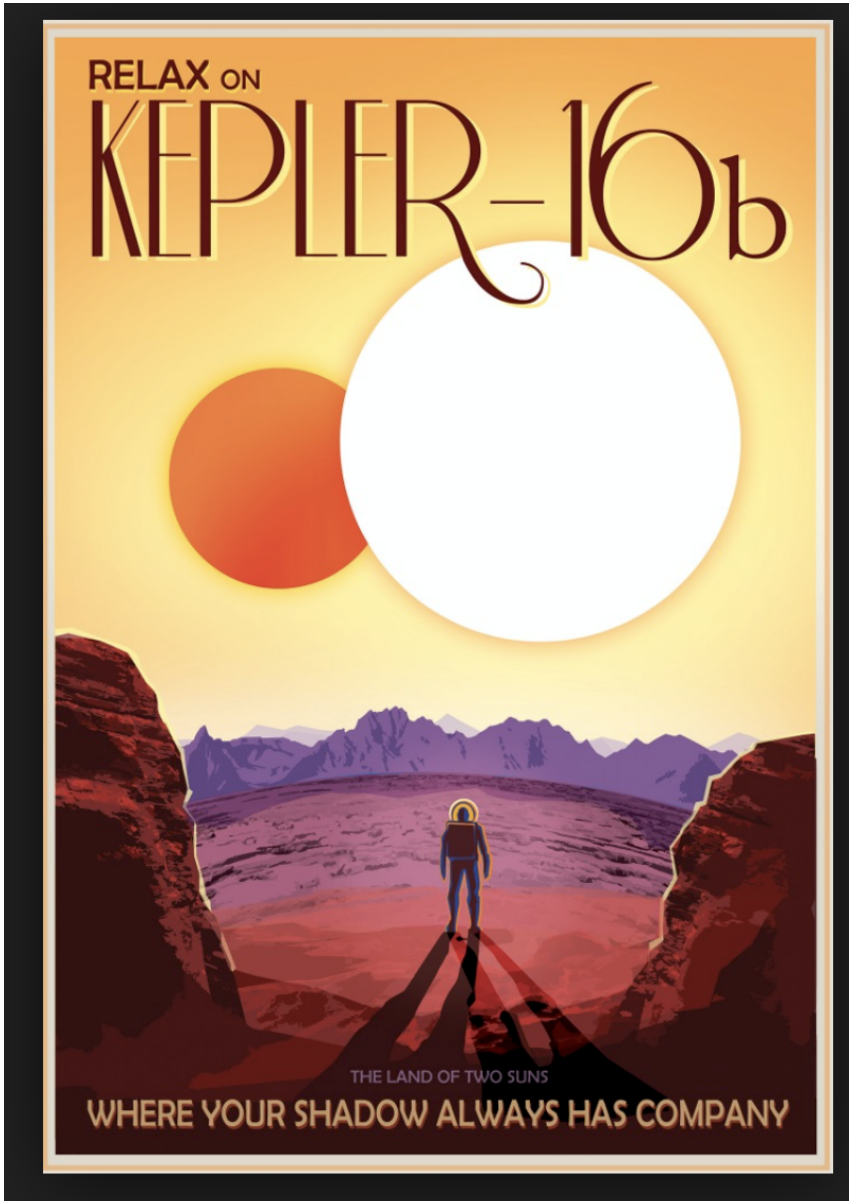
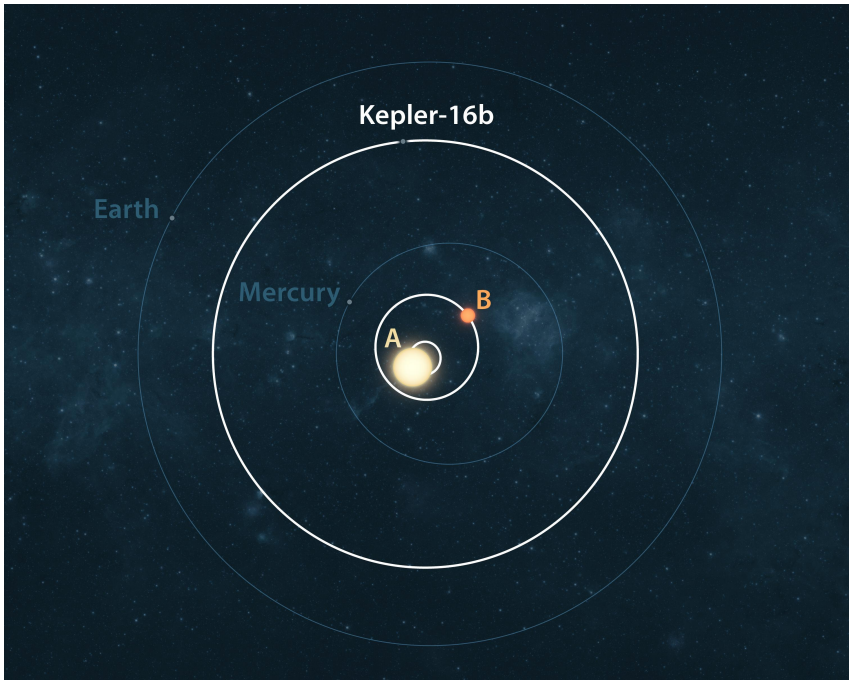


Image credit: https://www.skyimagelab.com/kepler-16-tour.html



- Kepler-16 contains the first discovered planet that orbits around a binary main-sequence star system (Doyle et al., 2011)

  (Known as a circumbinary planet)
- Stars are about 20% and 69% the mass of our Sun; the planet is comparable to Saturn (in mass and size) with about a 229-day orbit around the two stars
- Image credit: https://www.jpl.nasa.gov/spaceimages/details.php?id=PIA14727

Over 3800 exoplanets have been confirmed...

maybe you will discover the next one?

References

Adhikari, A. and DeNero, J. (2018), "Computational and Inferential Thinking," Gitbook.

Butler, R. P., Marcy, G. W., Williams, E., Hauser, H., and Shirts, P. (1997), "Three new "51 Pegasi-type" planets," *The Astrophysical Journal Letters*, 474, L115.

Doyle, L. R., Carter, J. A., Fabrycky, D. C., Slawson, R. W., Howell, S. B., Winn, J. N., Orosz, J. A., Prsa, A., Welsh, W. F., Quinn, S. N., et al. (2011), "Kepler-16: a transiting circumbinary planet," *Science*, 333, 1602–1606.

Kalas, P., Graham, J. R., Chiang, E., Fitzgerald, M. P., Clampin, M., Kite, E. S., Stapelfeldt, K., Marois, C., and Krist, J. (2008), "Optical images of an exosolar planet 25 light-years from Earth," *Science*, 322, 1345–1348.

Mayor, M. and Queloz, D. (1995), "A Jupiter-mass companion to a solar-type star," *Nature*, 378, 355.