

Now What?

Model Building, Science, and Analytics

Phil Henrickson, PhD
Senior Data Scientist
AE Business Solutions

Raise your hand if...

Raise your hand if...

You have taken an astronomy class.

Raise your hand if...

You have taken an astronomy class.

You have played a board game in the last year.

Raise your hand if...

You have taken an astronomy class.

You have played a board game in the last year.

You have watched **The Good Place**.

More to come on this later.

Now What?

Model Building, Science, and Analytics

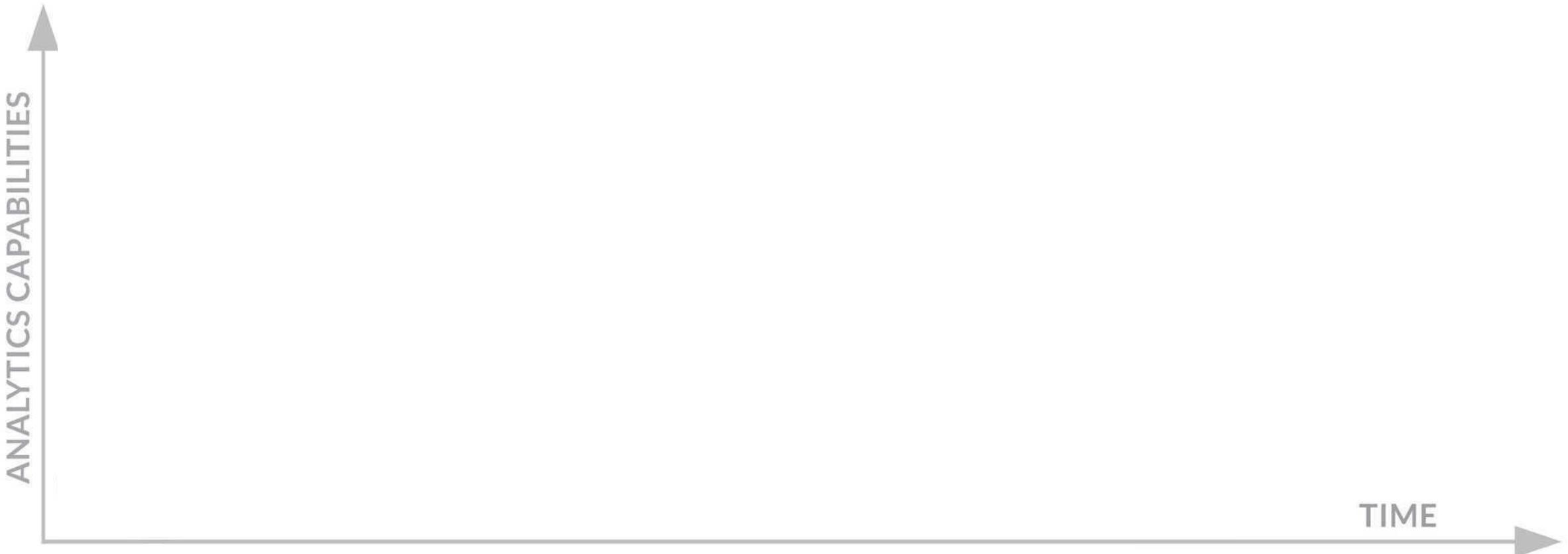
Phil Henrickson, PhD
Senior Data Scientist
AE Business Solutions

We've been seeing the same sentiment expressed amongst our clients more and more:

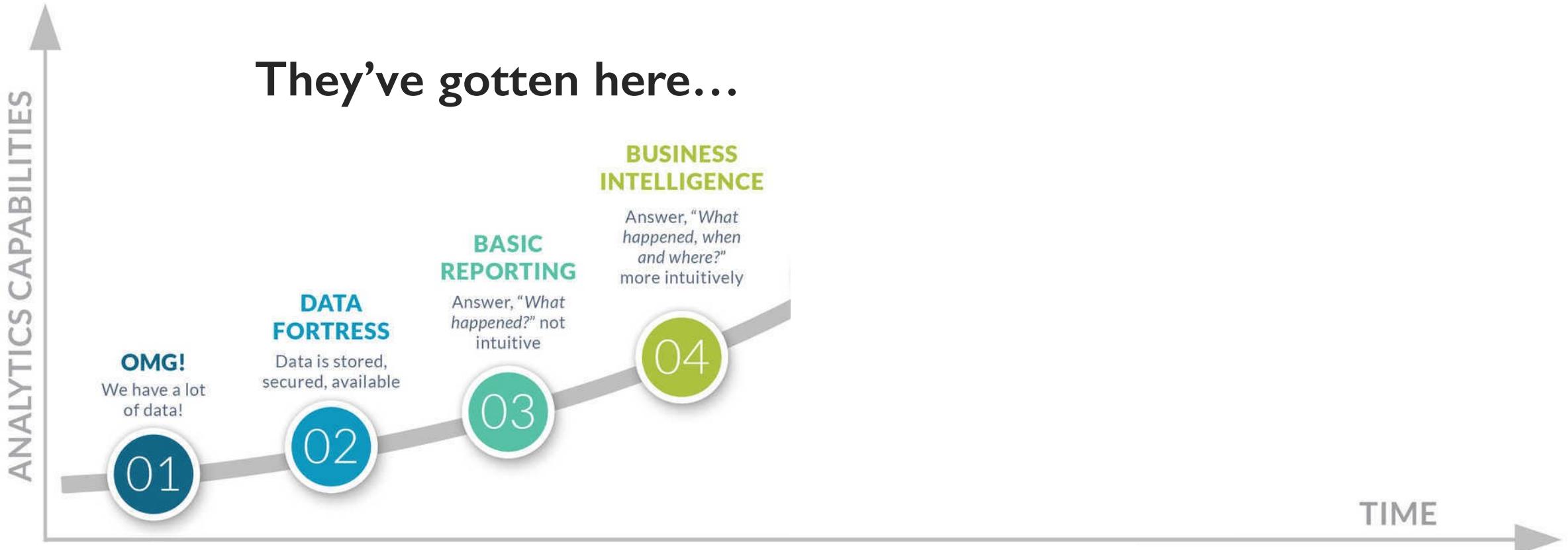
“We've modernized our data warehouse, we have all this data, we built all these dashboards...

Now what do we do?”

Most of our clients have discovered problems that dashboards and reports cannot solve.



Most of our clients have discovered problems that dashboards and reports cannot solve.



Data → Dashboard → ??? → #DataDrivenInsight



something was
supposed to happen
here



Data → Model → Dashboard → ??? → #AIDrivenInsight ™



*surely, something will
happen here*

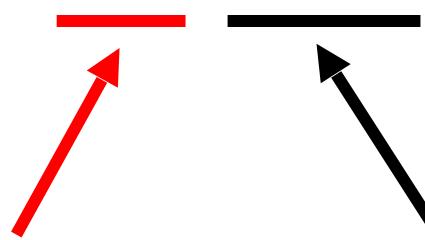
**The analytics community has been in a bit
of a frenzy over data science.**

The analytics community has been in a bit
of a frenzy over **data science**.



We need to
ignore this for a
second.

The analytics community has been in a bit
of a frenzy over **data science**.



We need to
ignore this for a
second.

We need to
spend more time
talking about
this.

science



We need to
spend more time
talking about
this.

science

We are all pretty comfortable
calling ourselves data people.

We hear the term ‘data-driven’
all the time.

science

**Should we strive to be
data driven?**

science

**Should we strive to be
data driven?**

I would argue no.

JUDEA PEARL

WINNER OF THE TURING AWARD

AND DANA MACKENZIE

THE
BOOK OF
WHY



THE NEW SCIENCE
OF CAUSE AND EFFECT

JUDEA PEARL

WINNER OF THE TURING AWARD

AND DANA MACKENZIE

THE
BOOK OF
WHY



THE NEW SCIENCE
OF CAUSE AND EFFECT

I hope to convince you that **data are profoundly dumb**.

No machine can derive explanations from raw data... Data can tell you that the people who took a medicine recovered faster than those who did not take it, but **they can't tell you why**.

JUDEA PEARL
WINNER OF THE TURING AWARD
AND DANA MACKENZIE

THE
BOOK OF
WHY

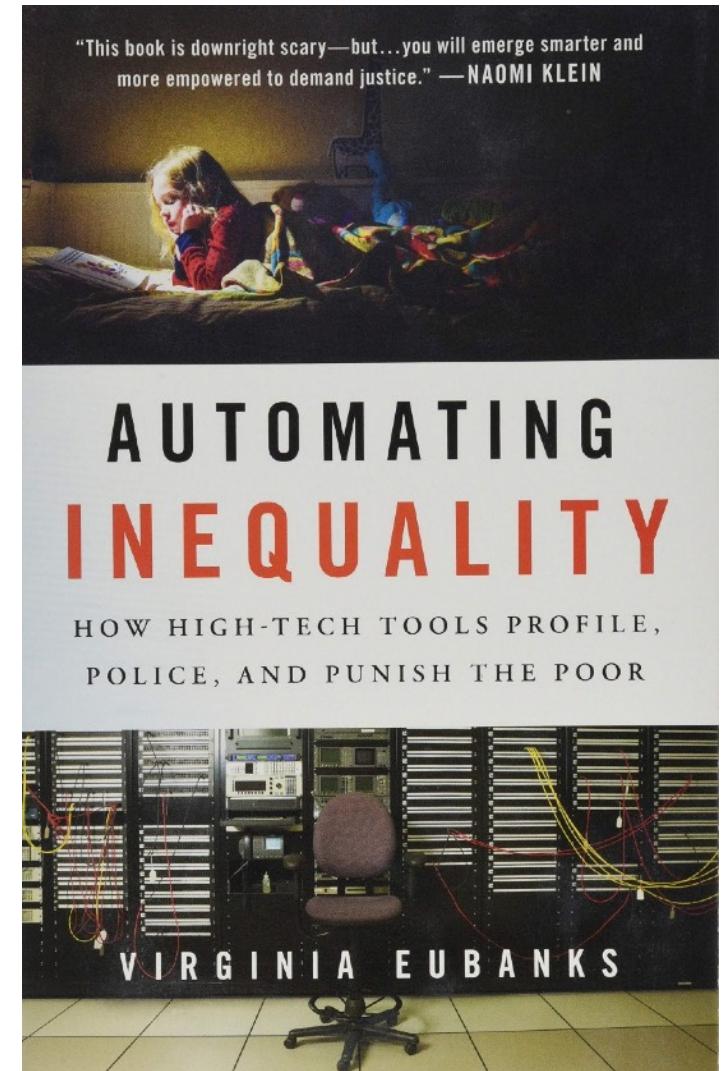
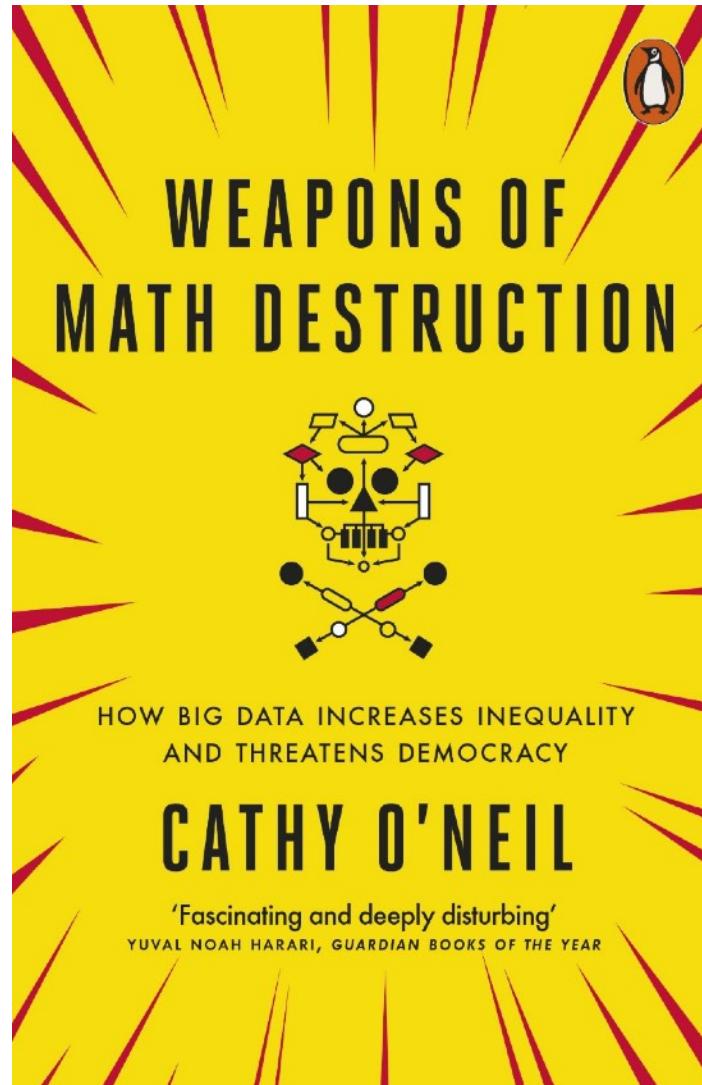
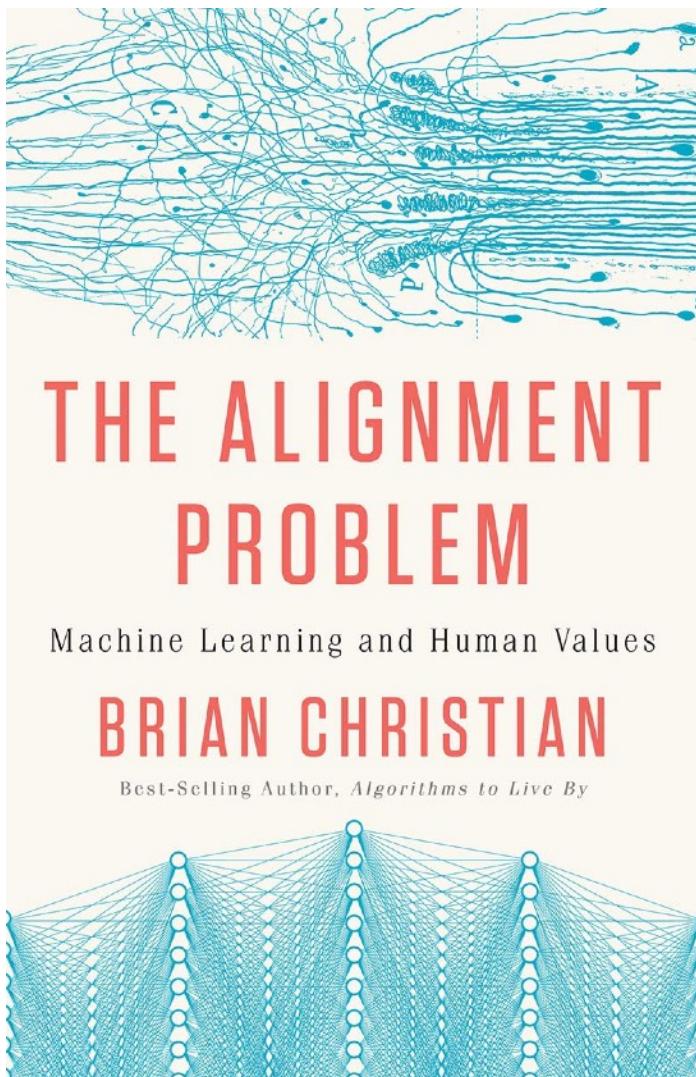


THE NEW SCIENCE
OF CAUSE AND EFFECT

Over and over again, in science and in business, we see situations where mere data aren't enough.

The hope... is that the data themselves will guide us to the right answers whenever causal questions come up.

If you'd like to read examples of data guiding to wrong answers:



JUDEA PEARL

WINNER OF THE TURING AWARD

AND DANA MACKENZIE

THE
BOOK OF
WHY



THE NEW SCIENCE
OF CAUSE AND EFFECT

If I could sum up the message of this book in one pithy phrase, it would be that **you are smarter than your data**.

Data do not understand causes and effects; **humans do**.

science

data science

**Data is an ingredient. It isn't a
recipe.**

**By itself, data offers no guarantee of
learning.**

data science

**Data is necessary for learning.
But it is not sufficient.**

data science

**Data is necessary for learning.
But it is not sufficient.**

**In order to learn from data, we
have to use a methodology.**

Data → Dashboard → ??? → #DataDrivenInsight



we were hoping to
learn something from
data



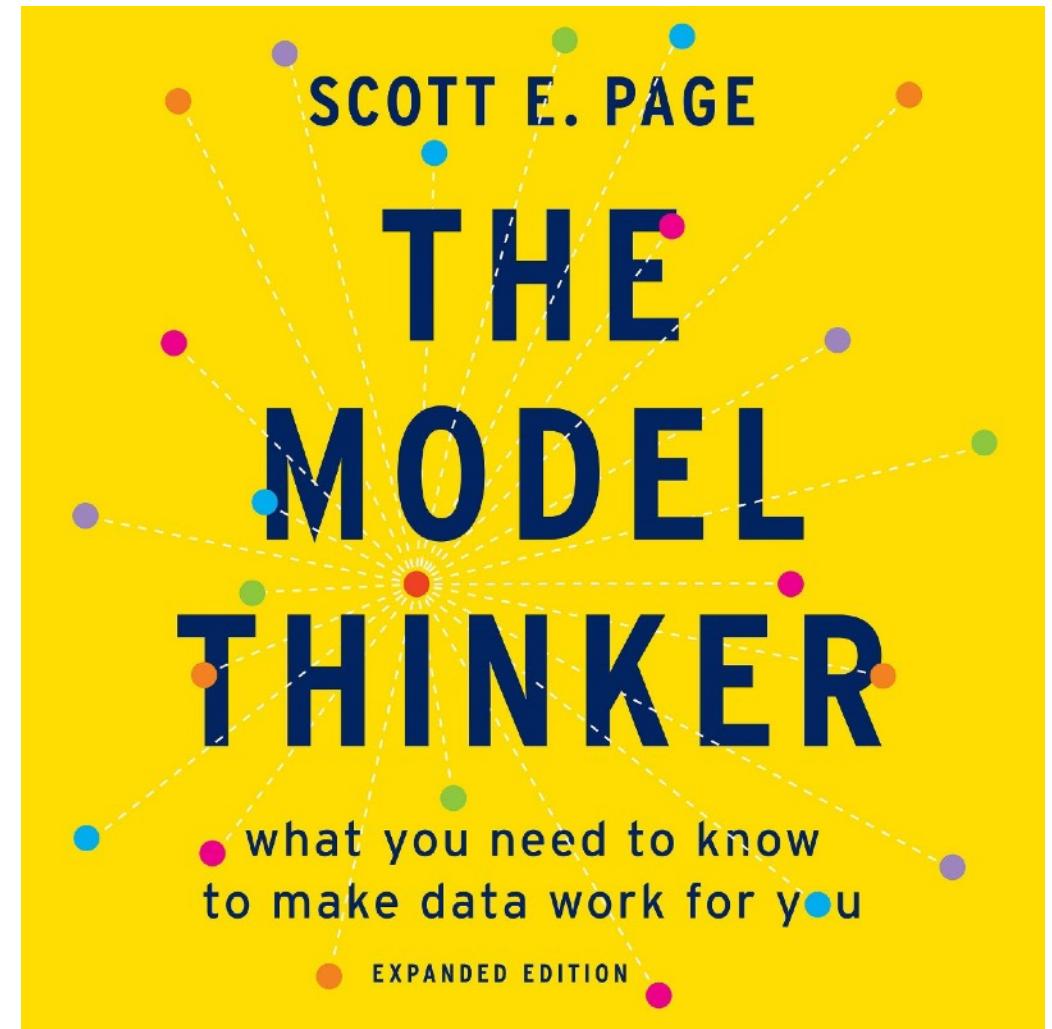
Data → Model → Dashboard → ??? → #AIDrivenInsight ™

data science

For all of the time we spend working with data, I don't think we spend enough time talking about **how we use it to learn.**

data science.

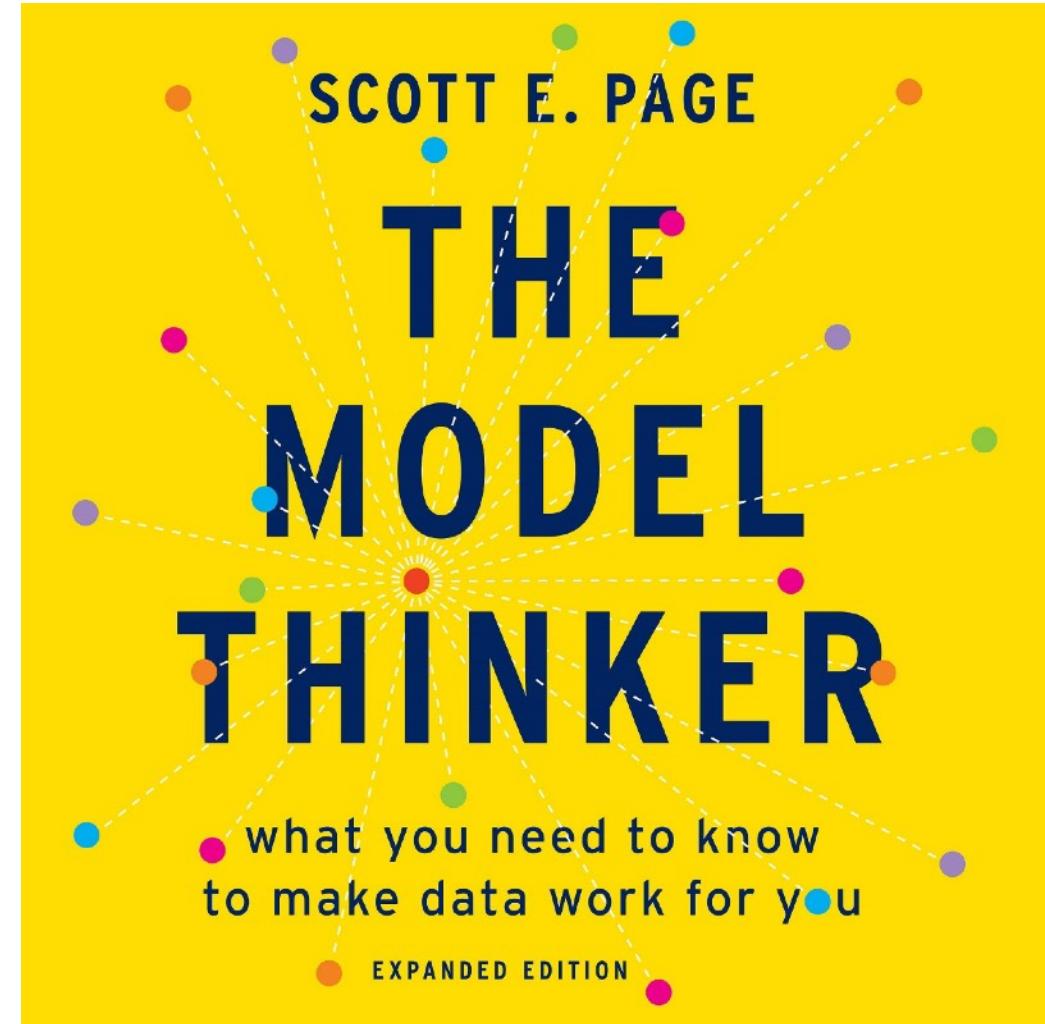
That brings us to the topic of the day: building models.



The rise of model thinking has [a simple explanation]: **models make us smarter.**

Without models, we have limited capacity to include data. With models, we clarify assumptions and think logically.

With models, **we think better.**



today

We need to talk about **models**.

today

We need to talk about why we build
models.

today

We need to talk about why we build
models.

We need to talk about how we build
models.

today

We need to talk about why we build
models.

We need to talk about how we build
models.

We need to talk about using models.

today

Part I

We need to talk about why we build models.

Part I & 2

We need to talk about how we build models.

Part 2

We need to talk about using models.

let's begin.

let's begin.

hold onto your butts.

1 Building Models of the Solar System

And Everything Else Along the Way

In a former life, I taught classes on
research methods, international conflict,
political violence, and civil war.

Regardless of the subject matter, I
started every semester with the same
lecture.

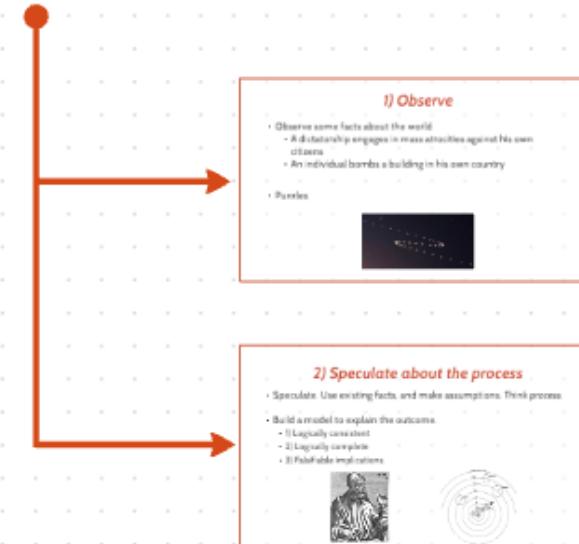
How to be a (Social) Scientist



A Model of the Model Building Process

- A model is a simplified representation of the real world.
- We create them by speculating about the processes that could have produced the observed facts.
- We evaluate models in terms of their ability to correctly predict facts, their generalizability, and their simplicity.

- 1) Observe
- 2) Speculate
- 3) Deduce implications
- 4) Test implications



Let's play a game!

The game board consists of a grid of boxes representing different stages of the model building process. The path starts at the bottom left, labeled 'Start', and moves through several stages: '1) Observe', '2) Speculate about the process', '3) Think of other implications', and '4) Test these implications'. Each stage contains a small image related to the process. The path then leads to a 'Finish' box at the top right. Along the way, there are several 'JAIL' boxes, which are labeled with 'You're out of jail' and 'Get back in jail' instructions. The board also features a 'GO' box and a 'LADY' box. The overall theme is a social science experiment or game.

How to be a (Social) Scientist

How to be a (Social) Scientist

People tend to diss on the social sciences,
mostly because they have a
misunderstanding of what social scientists
are up to.

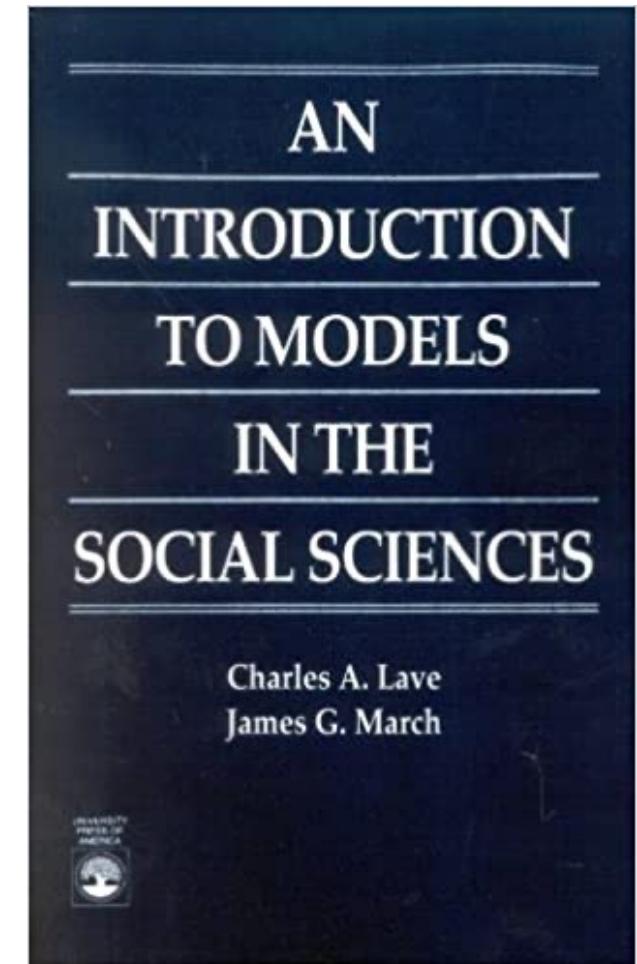
So, naturally, I immediately tried to win them
over using the coolest possible method:

How to be a (Social) Scientist

People tend to diss on the social sciences, mostly because they have a misunderstanding of what social scientists are up to.

So, naturally, I immediately tried to win them over using the coolest possible method:

a social sciences textbook from 1975



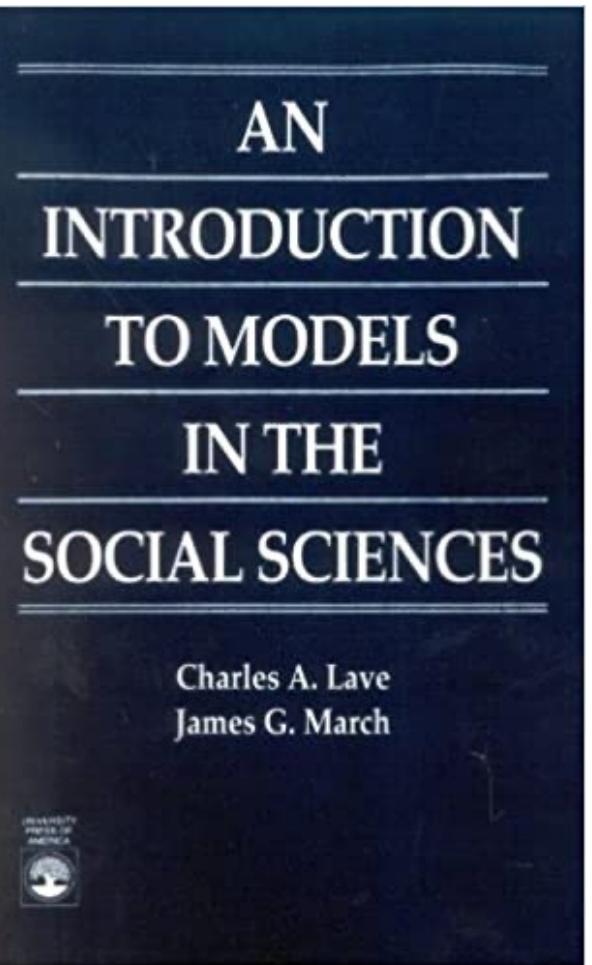
AN
INTRODUCTION
TO MODELS
IN THE
SOCIAL SCIENCES

Charles A. Lave
James G. March



*chapter
one*

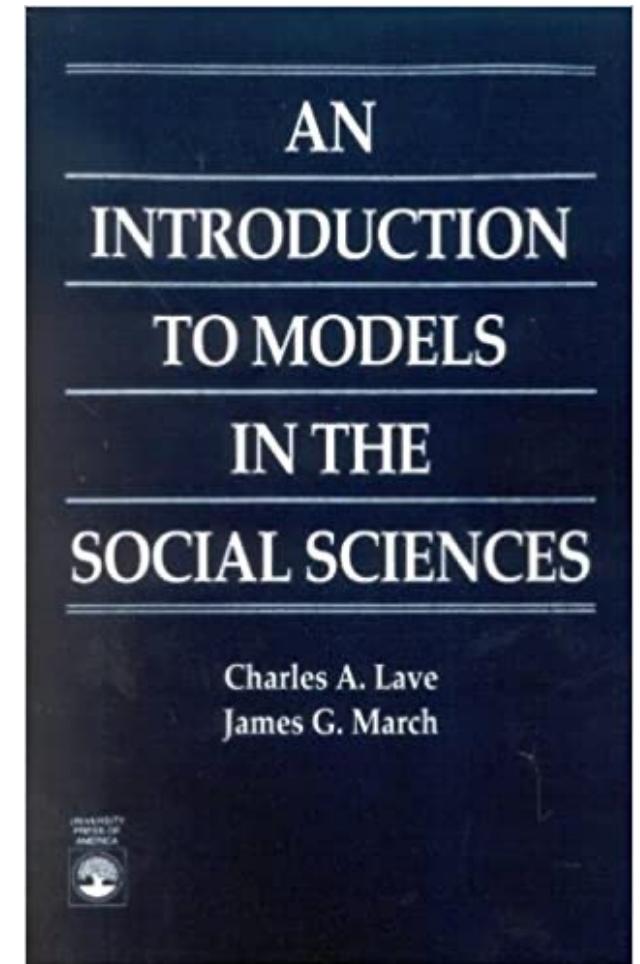
*what
we
are
up
to*



What is a model? How do you build one?
What makes a good model?

Speculative models are central to science, history, and literature. We are constantly forming partial interpretations of the world in order to live in it.

We think that an increase in the quality of speculation both in the social sciences and in everyday life would be good.



Lave and March are known for their
emphasis on **speculation**.

Lave and March are known for their emphasis on **speculation**.

They tell us to look at data as the end result of some process. Our goal is to **understand the process that produced the data**.

Lave and March are known for their emphasis on **speculation**.

They tell us to look at data as the end result of some process. Our goal is to **understand the process that produced the data**.

To do this, we need to speculate. We need to **stop and think**.

The best way to learn about building models is to do it.

This is the first example from their book.

Suppose we are interested in
understanding why some people are
friends and not others.

Suppose we are interested in understanding why some people are friends and not others.

We gather data on a college campus by asking residents of dormitories to give us a list of their friends.



Suppose we are interested in understanding why some people are friends and not others.

We gather data on a college campus by asking residents of dormitories to give us a list of their friends.

We notice a pattern: friends tend to live close to each other; they tend to have adjacent dormitory rooms.

Why might this be? Stop and think.

One possible explanation:

Campus housing lets students choose where to live in their dormitories. Students prefer to live by their friends. So, students ask campus housing to have friends as roommates or be put in adjacent rooms.

One possible explanation:

Campus housing lets students choose where to live in their dormitories. Students prefer to live by their friends. So, students ask campus housing to have friends as roommates or be put in adjacent rooms.

This is **our speculation about the process** that produced the data we observed.

That is, we have a basic model of the **prior state** of the world which may be able to account for what we observe in the **current state** of the world.

One possible explanation:

Campus housing lets students choose where to live in their dormitories. Students prefer to live by their friends. So, students ask campus housing to have friends as roommates or be put in adjacent rooms.

Is this a good model?

We have to ask: **if this model is correct, what else should we expect to observe?**

One possible explanation:

Campus housing lets students choose where to live in their dormitories. Students prefer to live by their friends. So, students ask campus housing to have friends as roommates or be put in adjacent rooms.

One possible explanation:

Campus housing lets students choose where to live in their dormitories. Students prefer to live by their friends. So, students ask campus housing to have friends as roommates or be put in adjacent rooms.

This model assumes that students had already known each prior to the start of the semester. It would predict different patterns of friendships between freshman and upperclassman dorms.

But, we notice this same pattern in a freshman dormitory.

One possible explanation:

Campus housing lets students choose where to live in their dormitories. Students prefer to live by their friends. So, students ask campus housing to have friends as roommates or be put in adjacent rooms.

Can our model still explain the pattern? Probably not.

So what do we? Try again. We stop and think.

Another possible explanation:

College students come from similar backgrounds and have a lot in common. Students who live near each other will frequently interact and discover what they have in common, leading to friendship..

Another possible explanation:

College students come from similar backgrounds and have a lot in common.

Students who live near each other will frequently interact and discover what they have in common, leading to friendship..

This would explain why we observe clusters of friends in all dormitories, including freshman.

Does that mean our model is correct?

Another possible explanation:

College students come from similar backgrounds and have a lot in common.

Students who live near each other will frequently interact and discover what they have in common, leading to friendship..

This would explain why we observe clusters of friends in all dormitories, including freshman.

Does that mean our model is correct?

No! We need to develop more implications, then gather data to put them to test.

How to be a (Social) Scientist

A model is a simplified representation of the world.

We create models by speculating about the processes that could have produced the data that we observe.

How to be a (Social) Scientist

A model is a simplified representation of the world.

We create models by speculating about the processes that could have produced the data that we observe.

We evaluate models in terms of their ability to predict what we observe, their ability to generalize, and their simplicity.

How to be a (Social) Scientist

A model of the model building process looks like this:

How to be a (Social) Scientist

A model of the model building process looks like this:

I) **We observe.** We notice a pattern or result that has occurred in the world.

How to be a (Social) Scientist

A model of the model building process looks like this:

- 1) **We observe.** We notice a pattern or result that has occurred in the world.
- 2) **We speculate.** We develop an explanation for the process that could have produced our observation.

How to be a (Social) Scientist

A model of the model building process looks like this:

- 1) **We observe.** We notice a pattern or result that has occurred in the world.
- 2) **We speculate.** We develop an explanation for the process that could have produced our observation.
- 3) **We develop implications.** We ask, if our speculation is correct, what else should we expect to observe?

How to be a (Social) Scientist

A model of the model building process looks like this:

- 1) **We observe.** We notice a pattern or result that has occurred in the world.
- 2) **We speculate.** We develop an explanation for the process that could have produced our observation.
- 3) **We develop implications.** We ask, if our speculation is correct, what else should we expect to observe?
- 4) **We test.** We look to see whether the other implications of our model are supported in the data.

How to be a (Social) Scientist

A model of the model building process looks like this:

- 1) **We observe.**
- 2) **We speculate.**
- 3) **We develop implications.**
- 4) **We test.**

This model building process is applicable
far beyond the social sciences.



Mars at
day 25

Mars at
day 3

Mars at
day 2

Mars at
day 8

Mars at
day 40

Why does Mars move backwards in
the nighttime sky?

Model (Speculation)

**The Earth is at the center of the solar system.
The heavens are in perfect harmony and
objects orbit the Earth in circles.**

Model (Speculation)

Implication

The Earth is at the center of the solar system.
The heavens are in perfect harmony and
objects orbit the Earth in circles.

If that's the case, planets shouldn't move
backwards.

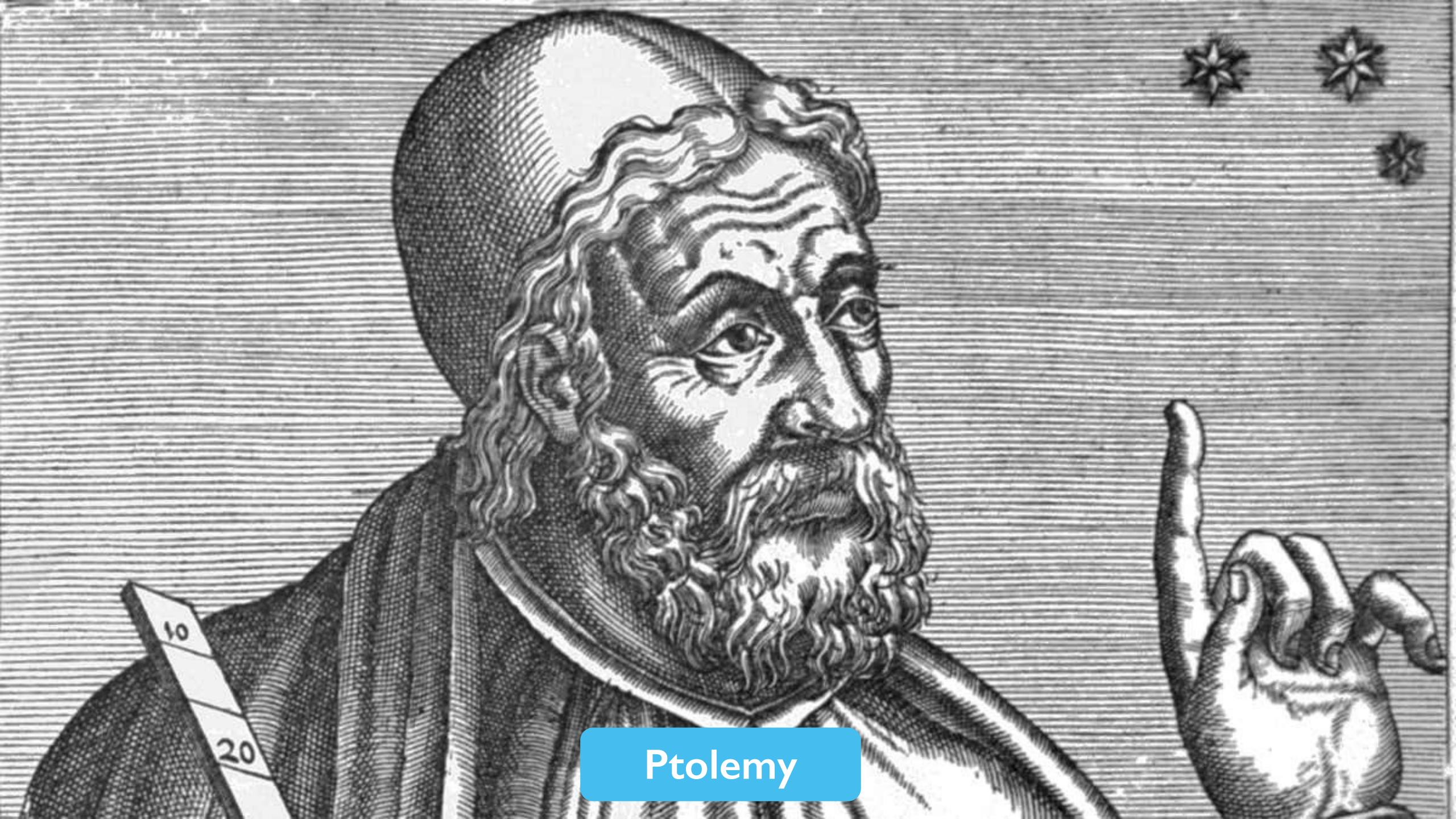
Model (Speculation)

The Earth is at the center of the solar system.
The heavens are in perfect harmony and
objects orbit the Earth in circles.

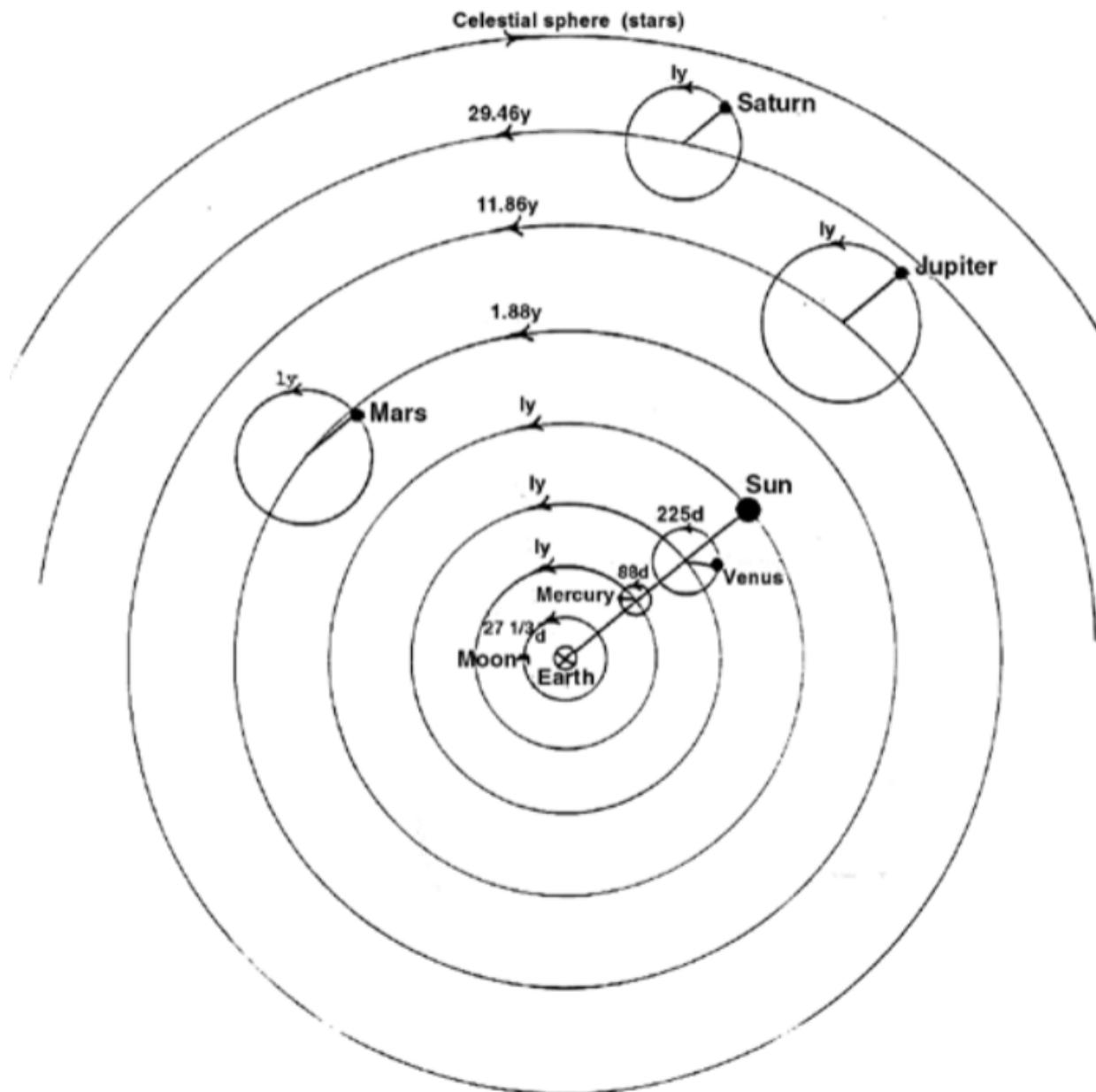
Implication

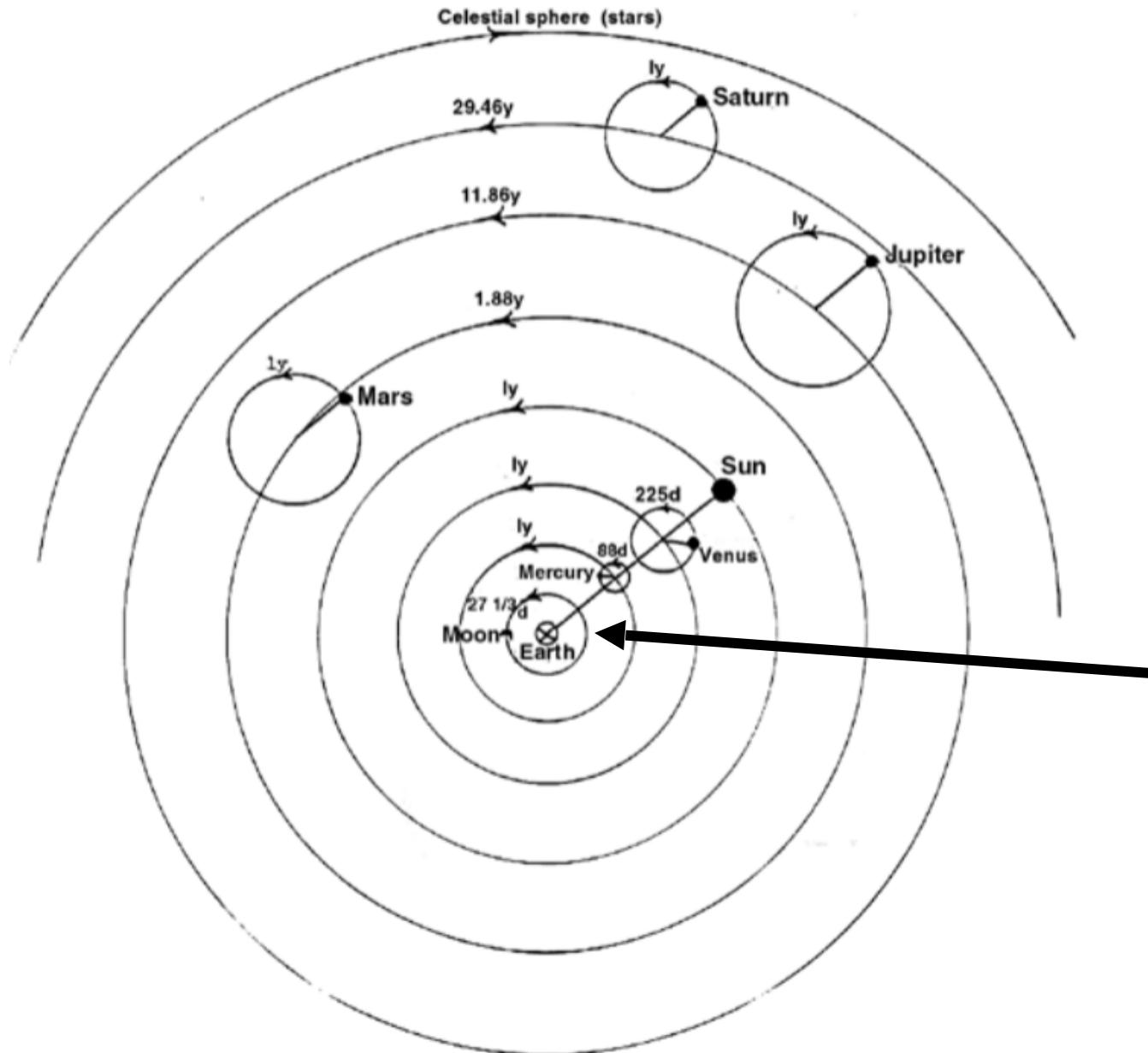
If that's the case, planets shouldn't move
backwards.

But they do.



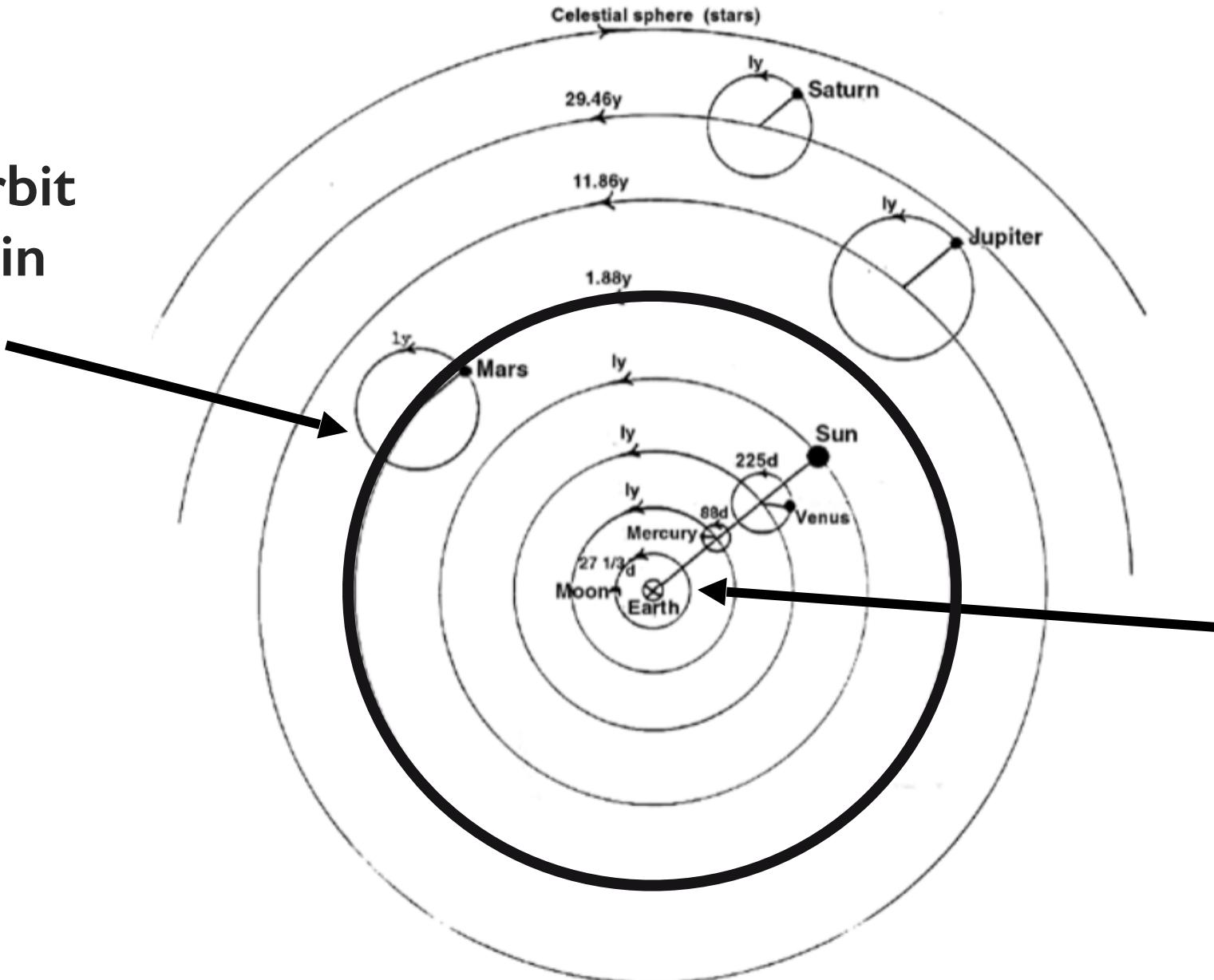
Ptolemy





**Earth at the
center**

**Planets orbit
the earth in
circles..**

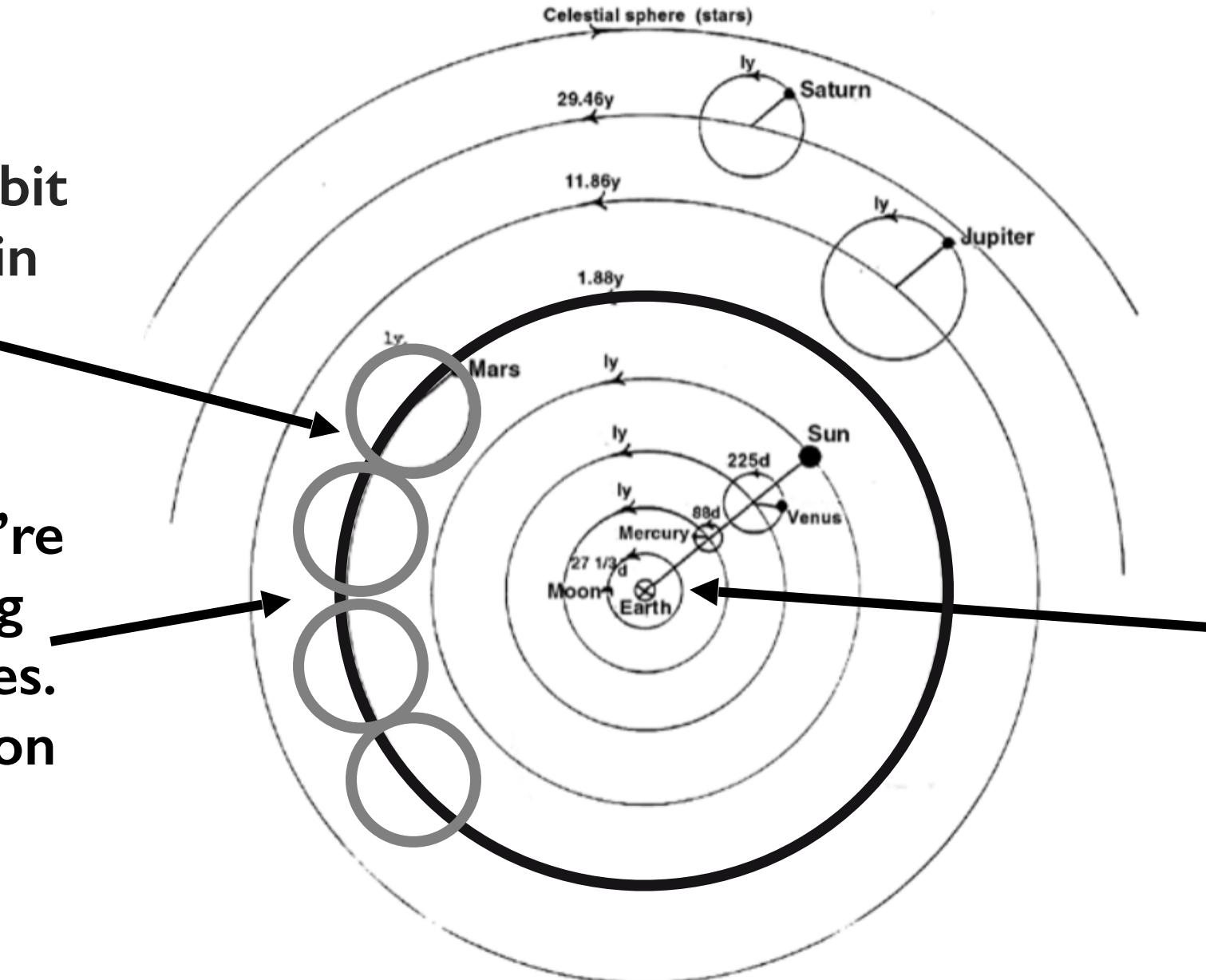


**Earth at the
center**

**Planets orbit
the earth in
circles..**

**...but they're
also making
other circles.
It's circles on
circles!**

**Earth at the
center**



If Ptolemy's model is correct, what should we observe?

If Ptolemy's model is correct, what should we observe?

It should be able to predict the movement of the planets.

If Ptolemy's model is correct, what should we observe?

It should be able to predict the movement of the planets.

And it does!

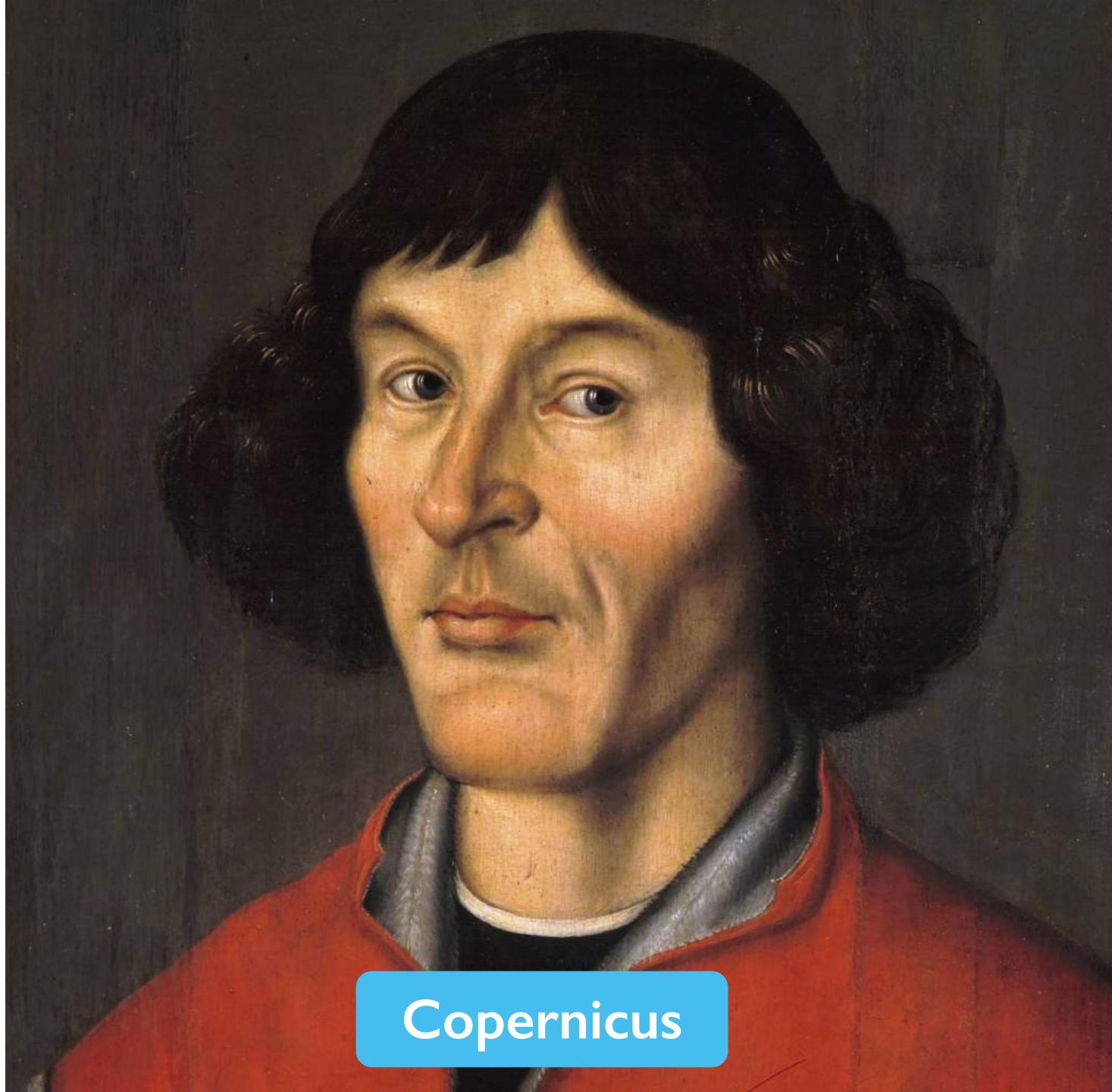
If Ptolemy's model is correct, what should we observe?

It should be able to predict the movement of the planets.

And it does! But not perfectly.

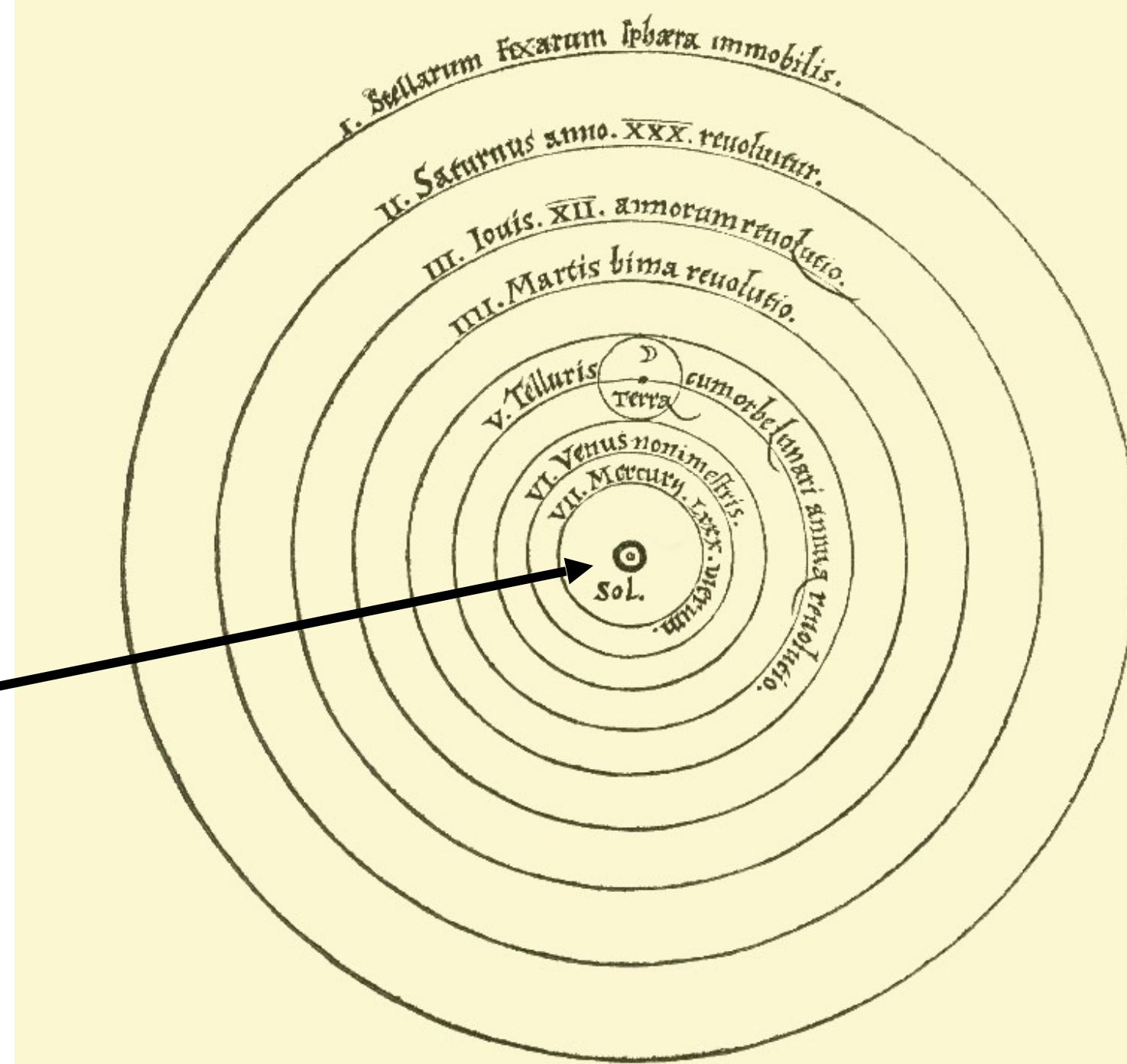
In the ensuing centuries, more observation led to more data collection which led to more inaccuracies with the Ptolemaic model.

New models for the solar system were proposed.

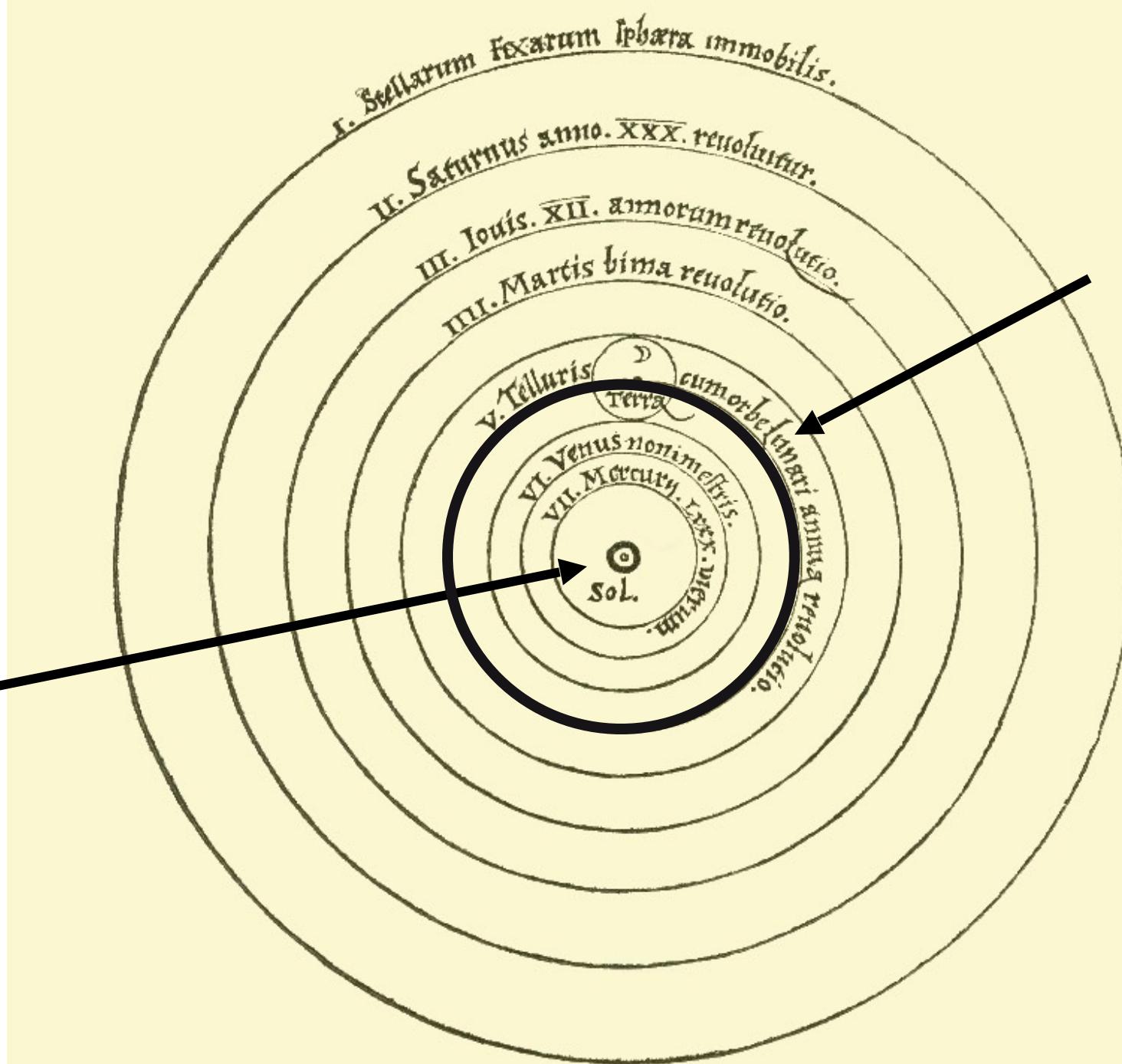


Copernicus

Sun at the center

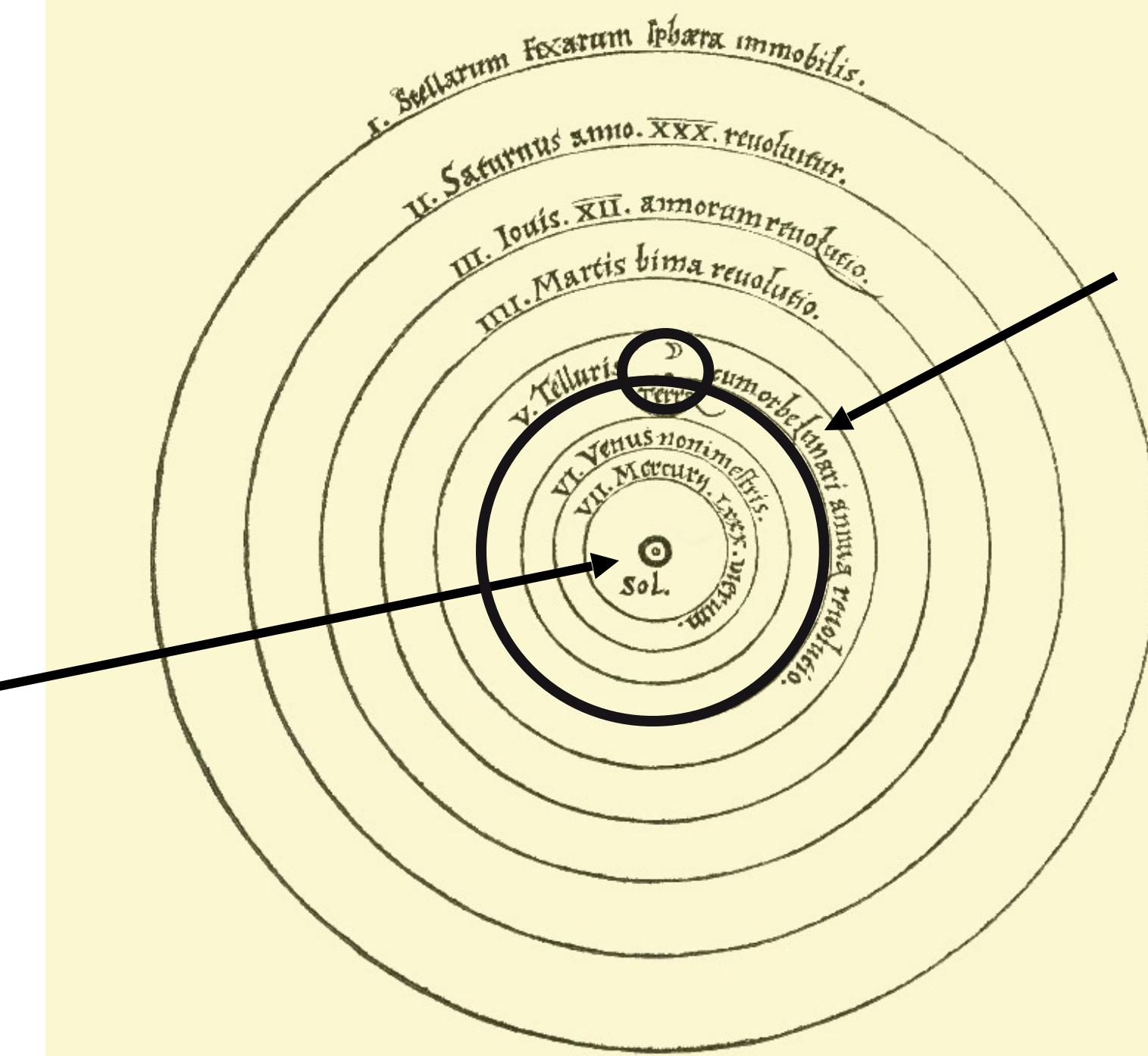


Sun at the center



The planets,
including
Earth, orbit
the sun in
circles

Sun at the center



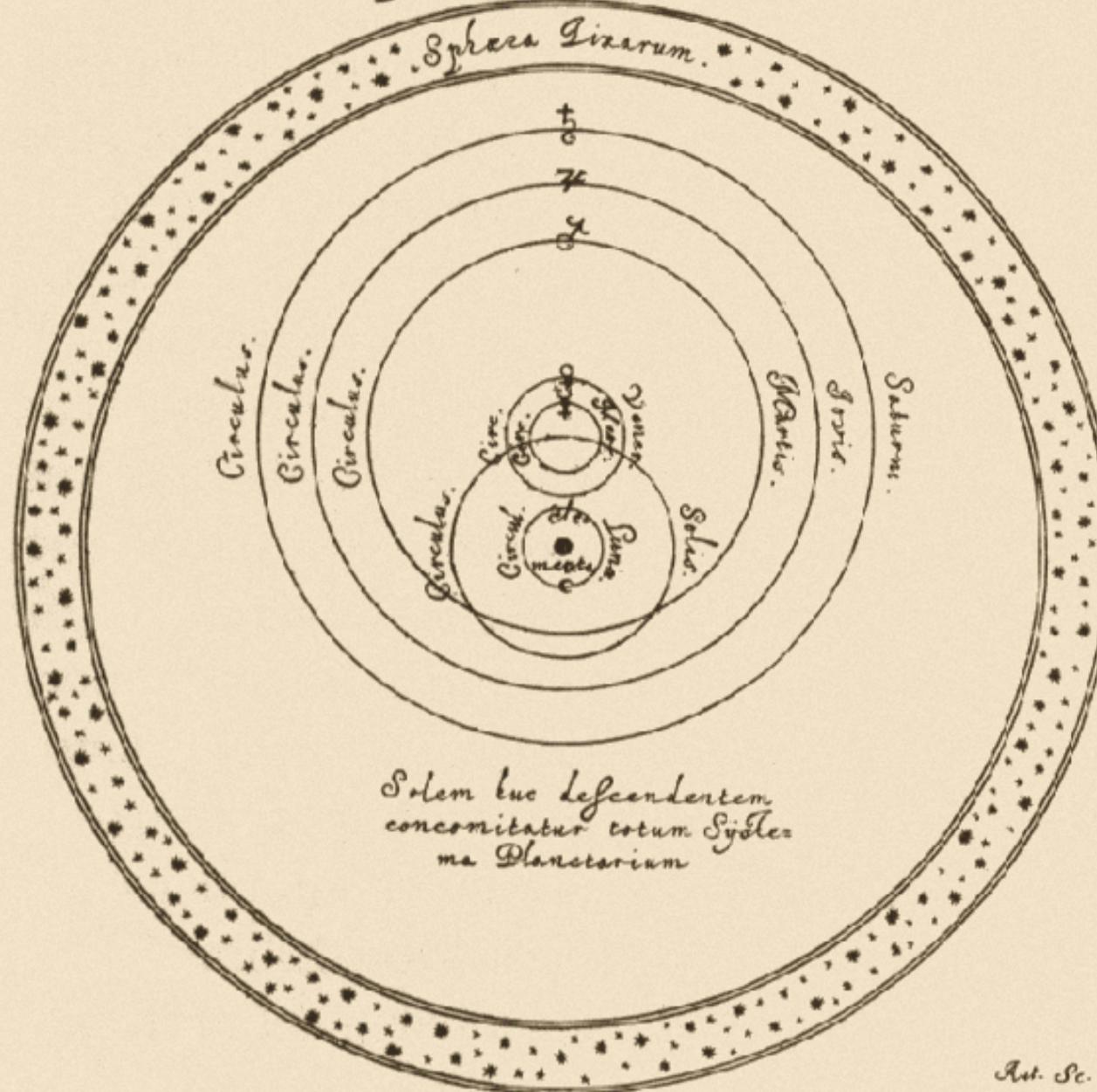
The planets,
including
Earth, orbit
the sun in
circles

Except for
the moon,
which orbits
Earth.

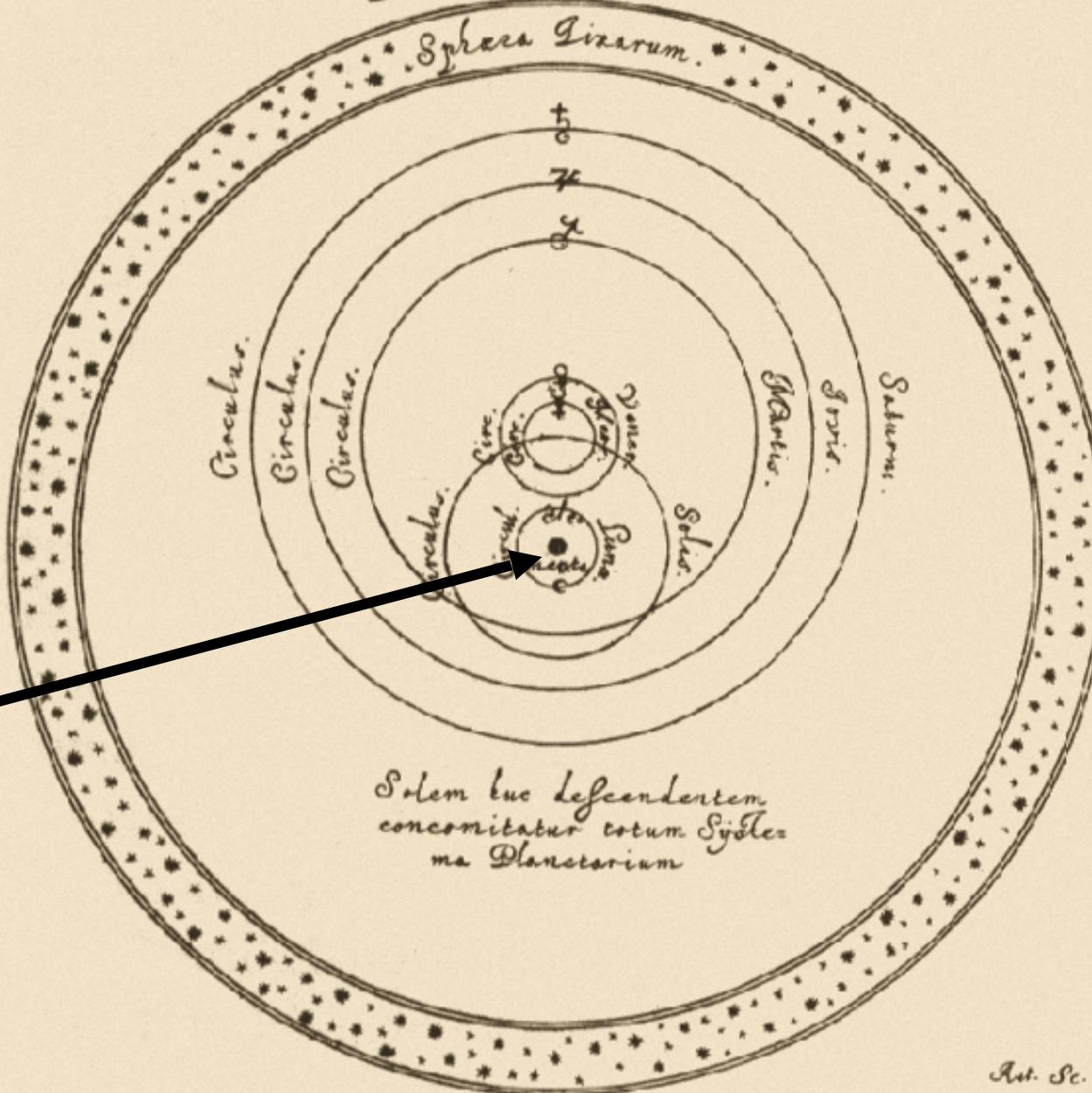


Tycho

Hypothesis Tychoonica.

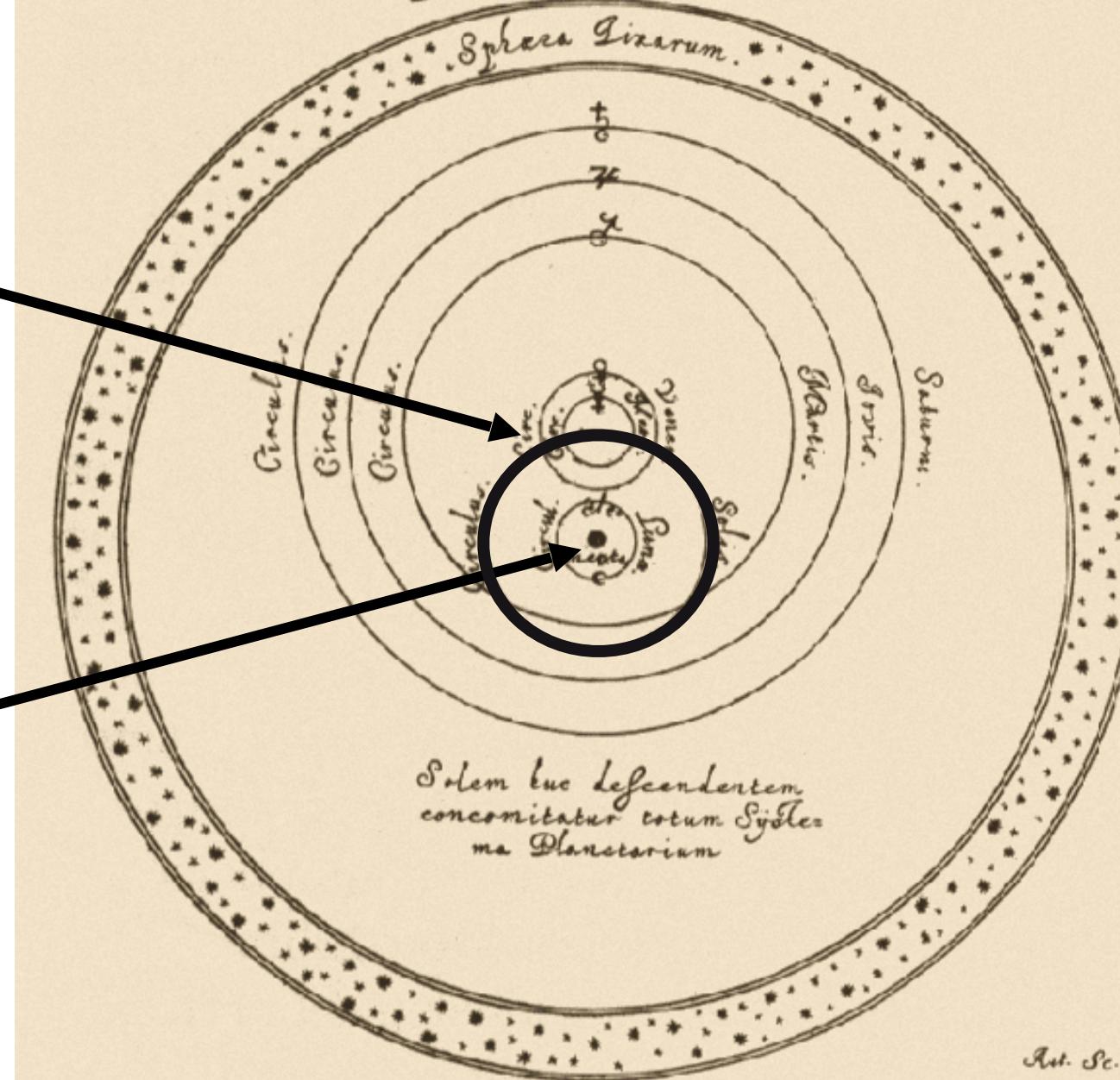


Hypothesis Tychoonica.



Earth at the
center

Hypothesis Tychonica.



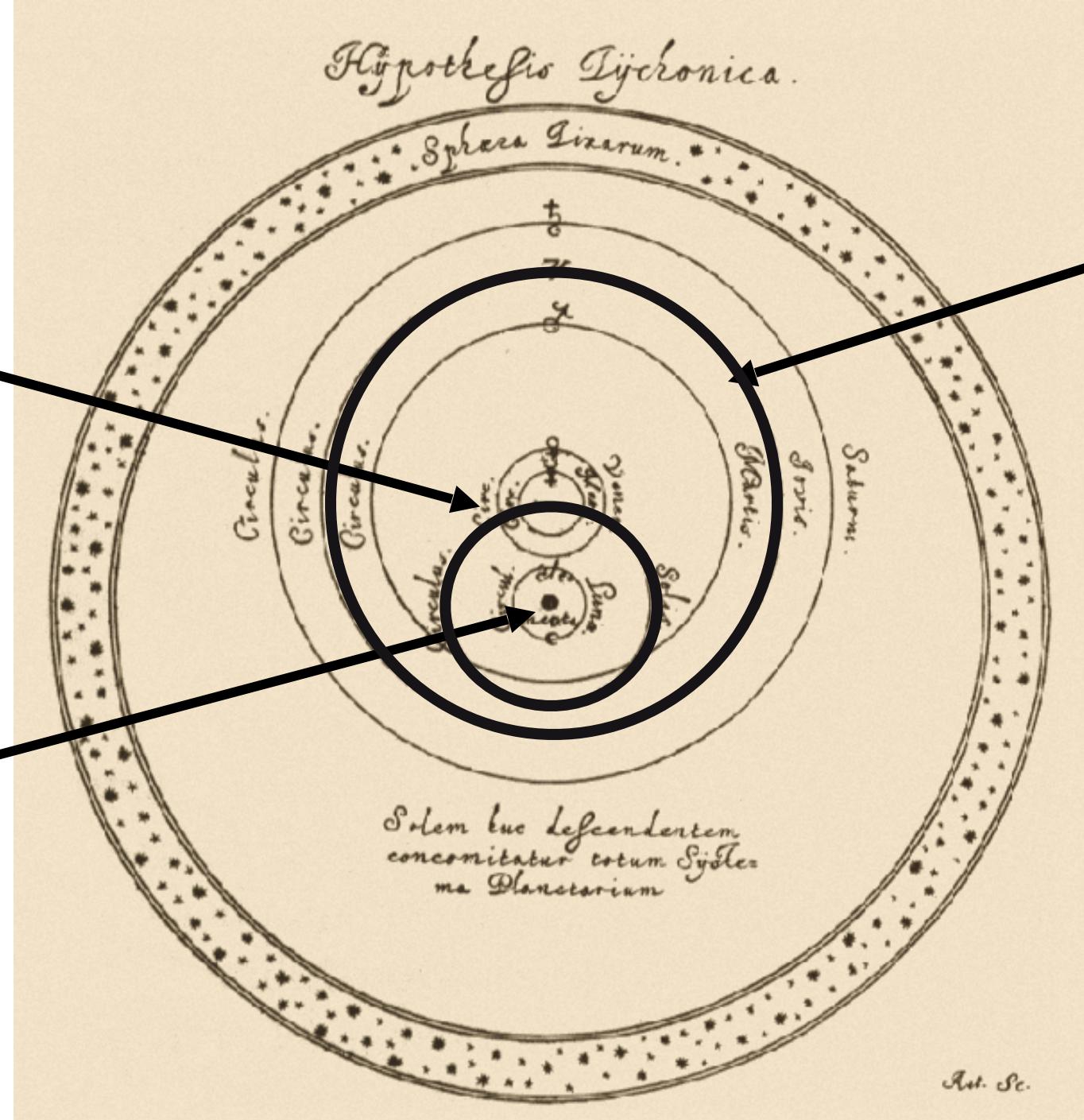
**Sun and
moon orbit
the Earth**

**Earth at the
center**

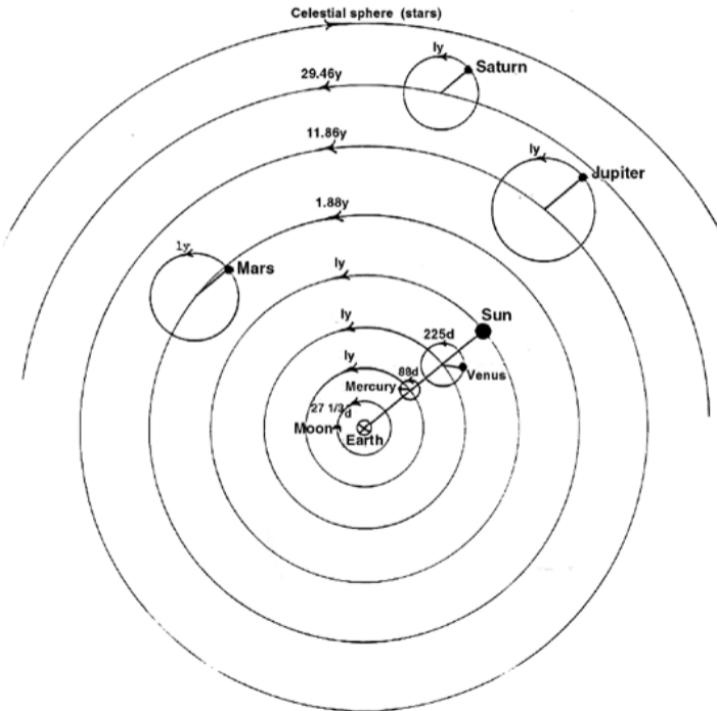
**Sun and
moon orbit
the Earth**

**Earth at the
center**

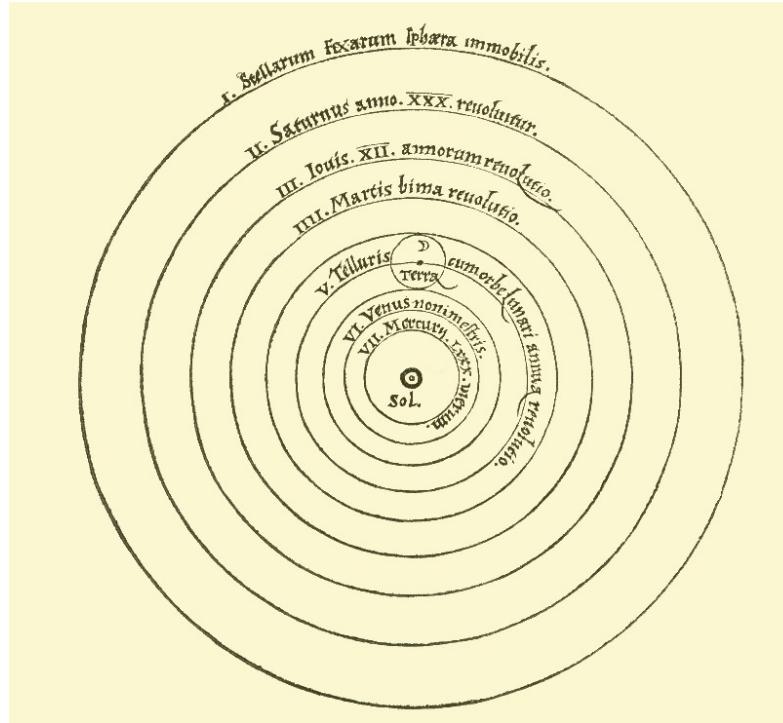
**The other
planets orbit
the Sun**



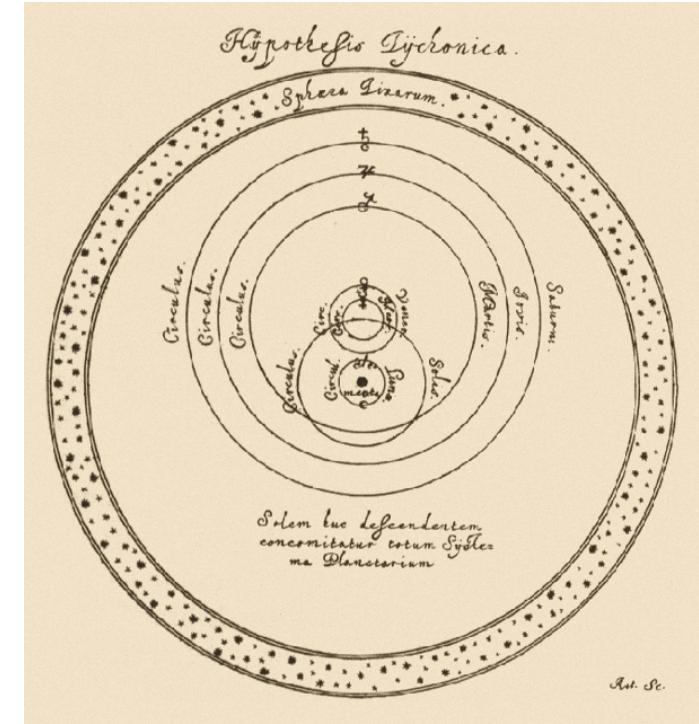
Why does Mars move backwards in the nighttime sky?



Ptolemy



Copernicus



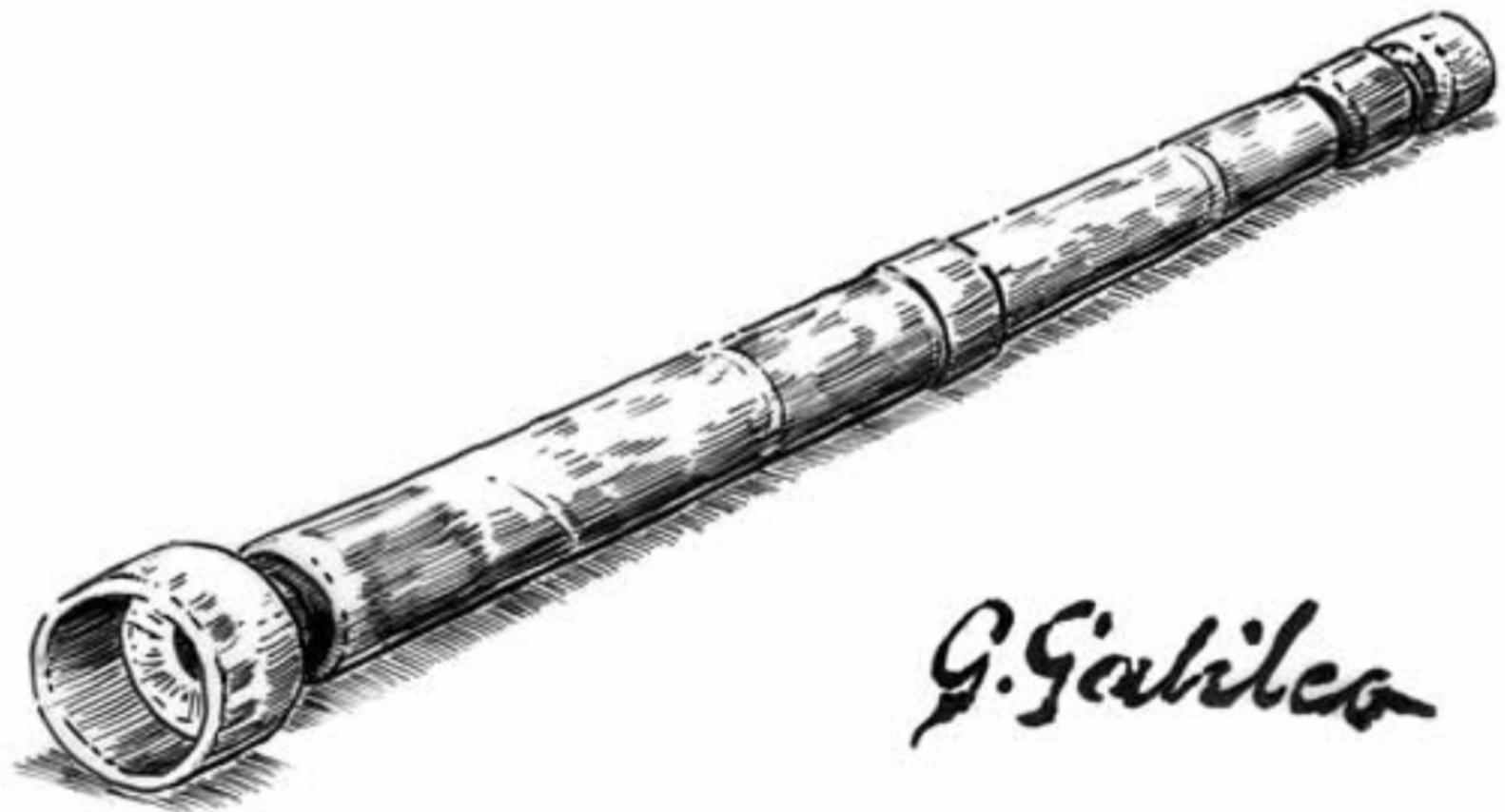
Tycho

How do we determine which model is
‘correct’?

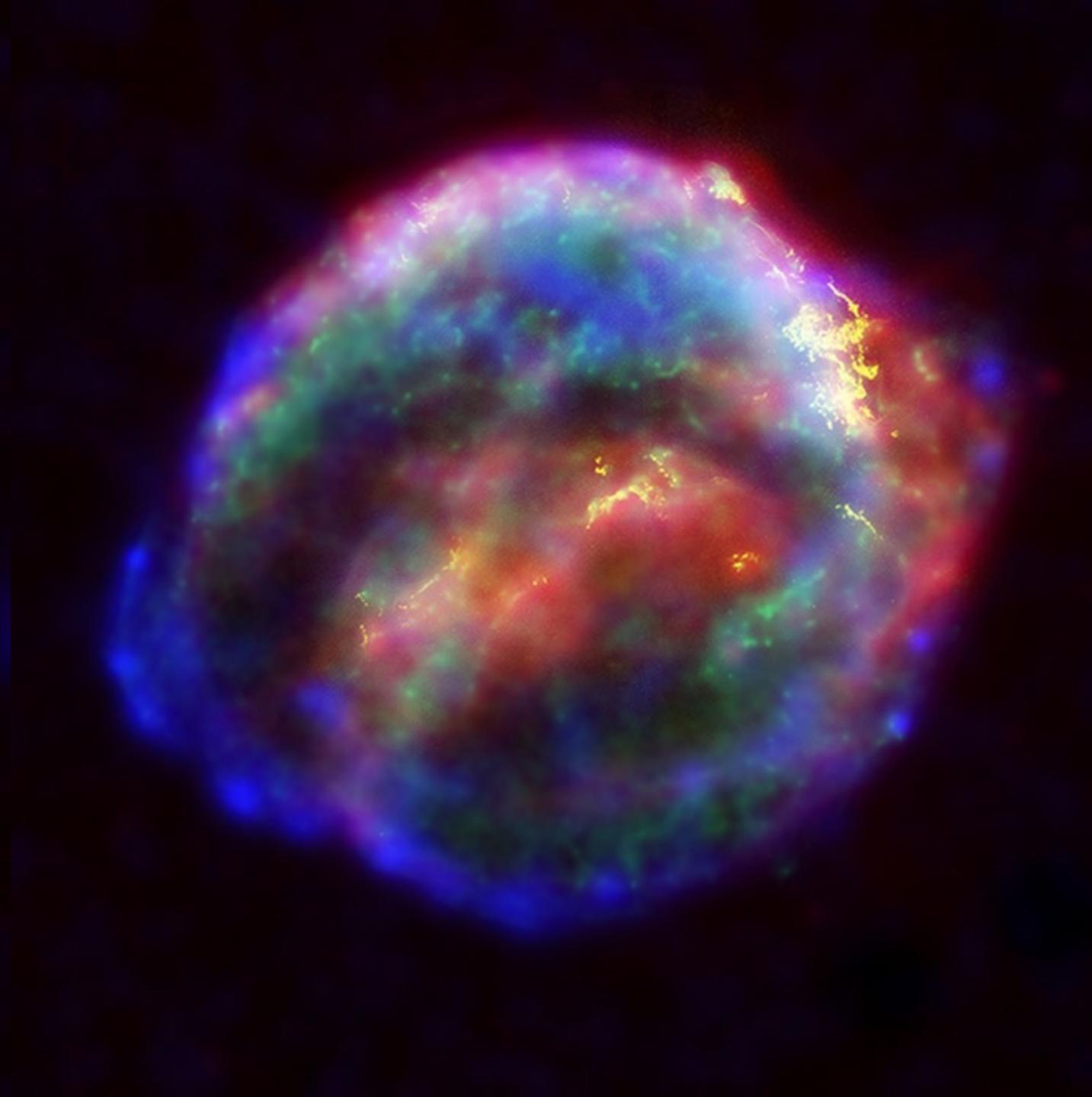
How do we determine which model is
‘correct’?

We think about implications. We gather
data.

New methods of data collection challenged
implications of Ptolemy's model.



G. Gabilca

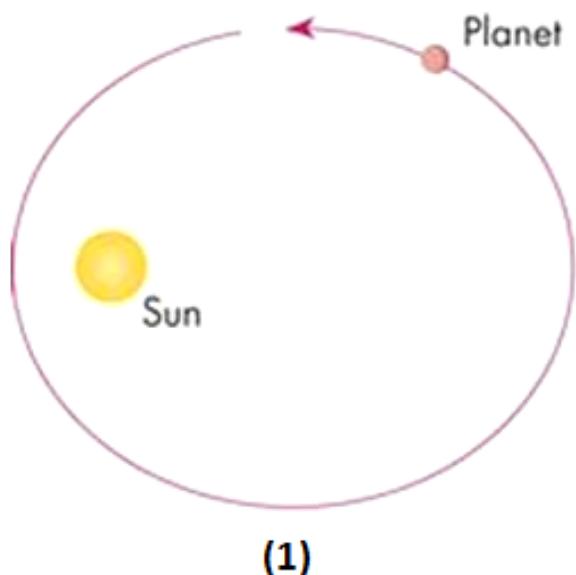


Kepler's
Supernova,
1604

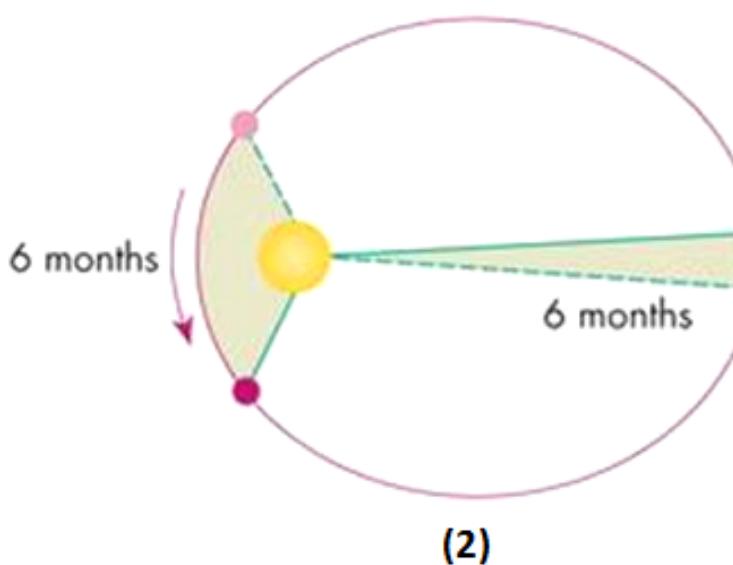


Kepler

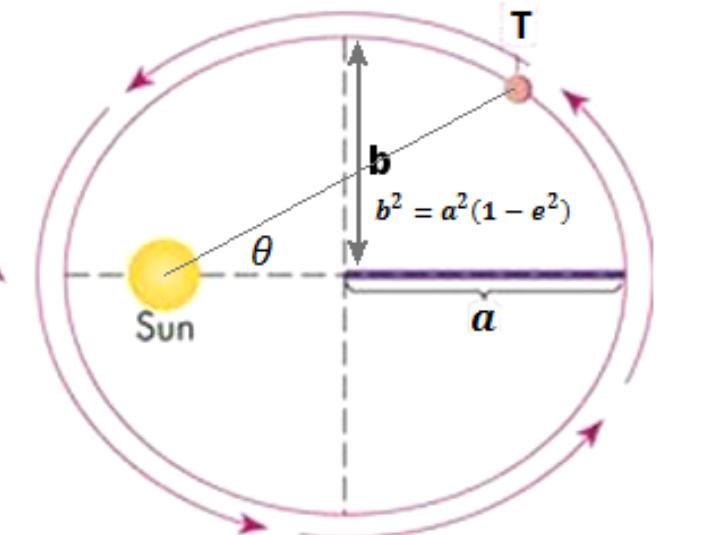
Kepler's Laws of Planetary Motion



The orbits are ellipses



Equal areas in equal time



$T^2 \propto a^3$ T = time to complete orbit
 a = semi-major axis

Then some guy named Isaac Newton
came along to explain why Kepler's
model was evidence of a more general
theory...

Newton's Principia in 1687

PHILOSOPHIAE NATURALIS PRINCIPIA MATHEMATICA.

Autore J S. NEWTON, Trin. Coll. Cantab. Soc. Mathefeos
Professore Lucasiano, & Societatis Regalis Sodali.

IMPRIMATUR
S. PEPYS, Reg. Soc. PRÆSES.
Julii 5. 1686.

LONDINI,

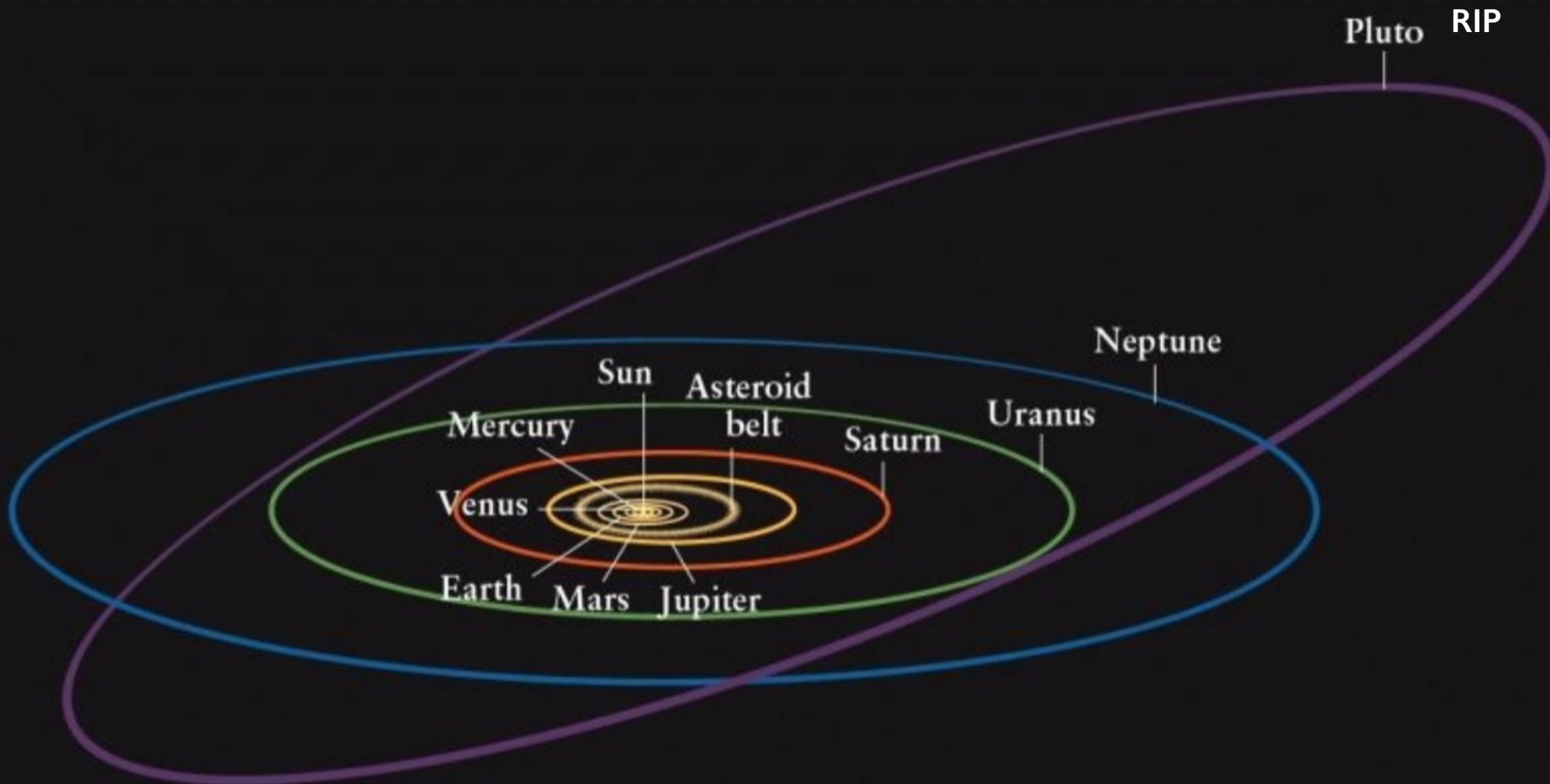
Jussu Societatis Regiae ac Typis Josephi Streater. Prostat apud
plures Bibliopolas. Anno MDCLXXXVII.

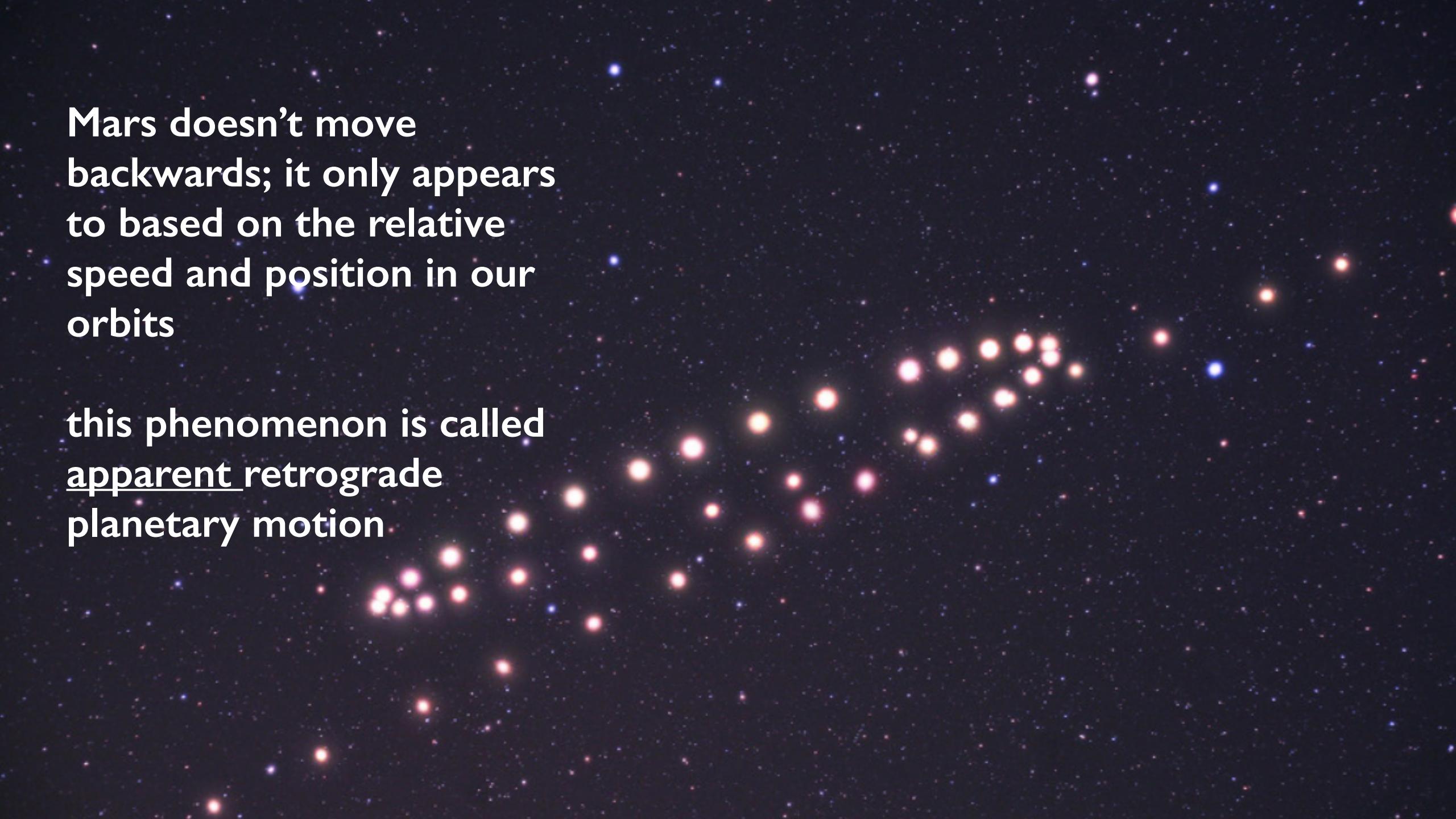
Newton's Laws of Motion

$$\sum \mathbf{F} = 0 \Rightarrow \frac{dv}{dt} = 0$$

$$F = ma$$

$$F_A = -F_B$$





Mars doesn't move
backwards; it only appears
to based on the relative
speed and position in our
orbits

this phenomenon is called
apparent retrograde
planetary motion

A model of the model building process:

- 1) **We observe.** We notice a pattern or result that has occurred in the world.
- 2) **We speculate.** We develop explanation for the process that could have produced our observation.
- 3) **We develop implications.** We ask, if our speculation is correct, what else should we expect to observe?
- 4) **We test.** We look to see whether the other implications of our model are supported in the data.

A model of the model building process:

-
- Data is used here.
- 1) **We observe.** We notice a pattern or result that has occurred in the world.
 - 2) **We speculate.** We develop explanation for the process that could have produced our observation.
 - 3) **We develop implications.** We ask, if our speculation is correct, what else should we expect to observe?
 - 4) **We test.** We look to see whether the other implications of our model are supported in the data.

A model of the model building process:

Data is used here.

1) **We observe.** We notice a pattern or result that has occurred in the world.

2) **We speculate.** We develop explanation for the process that could have produced our observation.

3) **We develop implications.** We ask, if our speculation is correct, what else should we expect to observe?

4) **We test.** We look to see whether the other implications of our model are supported in the data.

The science,
**how we learn
from data,** is
here.

Data does not, by itself, lead to learning.

**Data is the means of testing the implications
of our speculative models.**

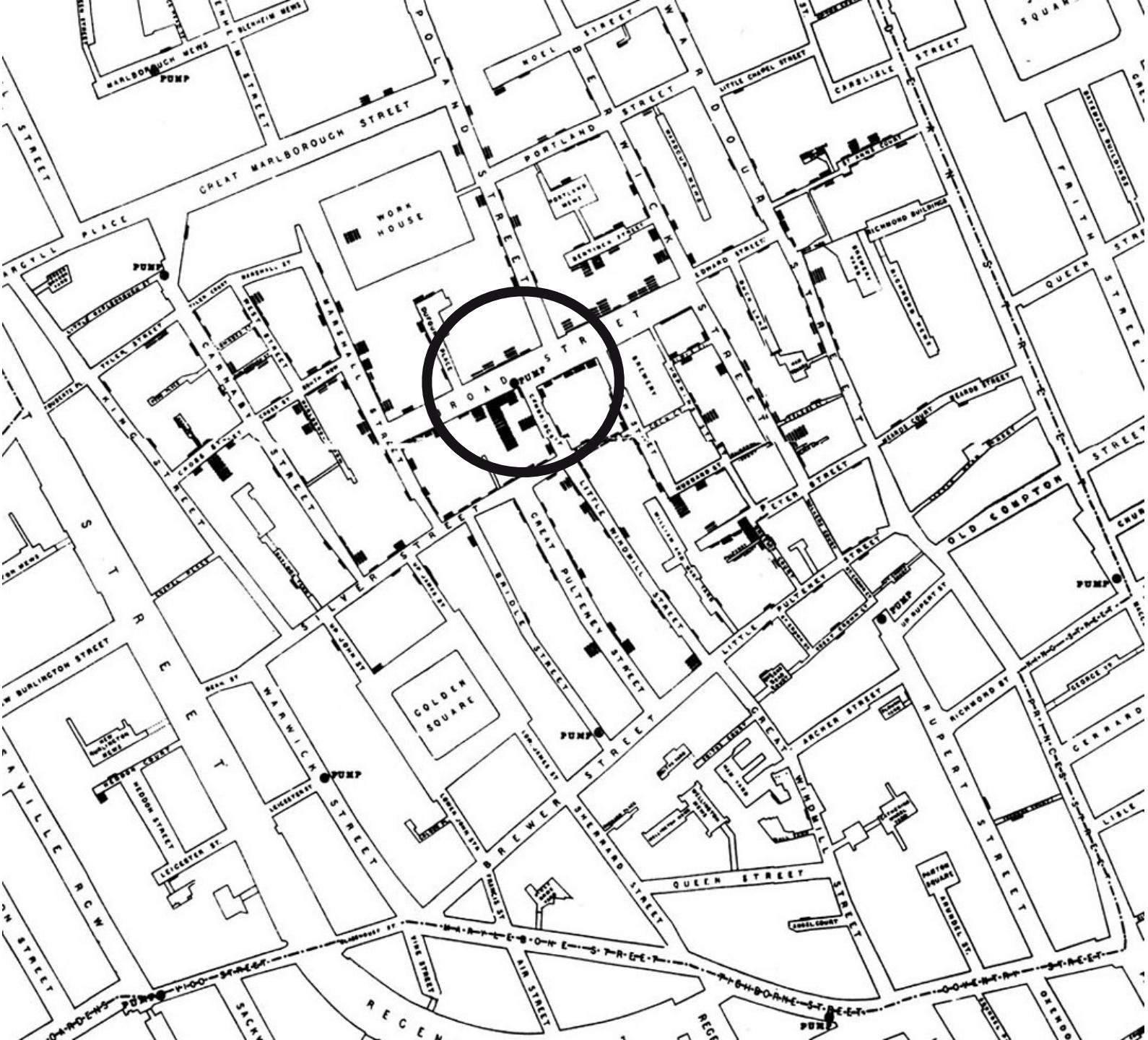
Example:

**John Snow and the London
Cholera Outbreak of 1854**

John Snow's map of cholera cases linked to the outbreak

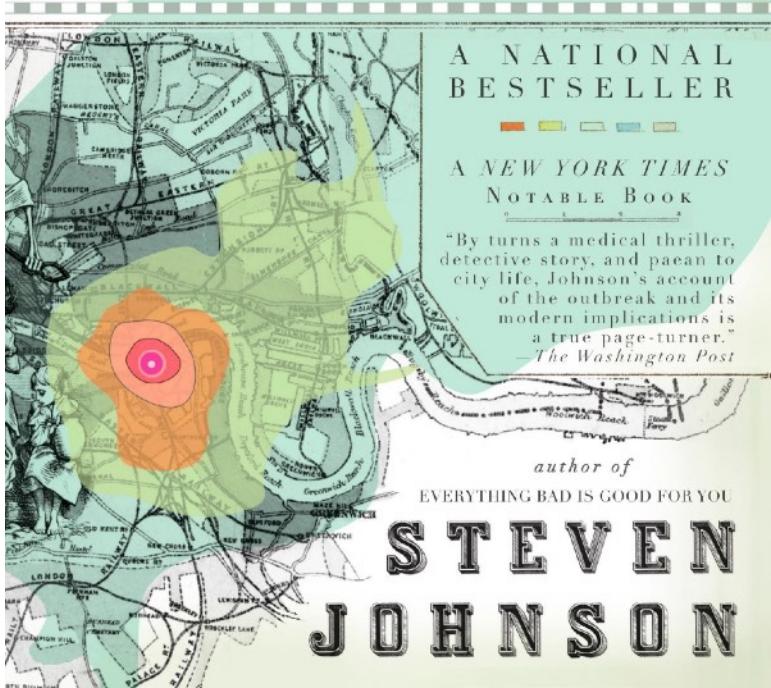


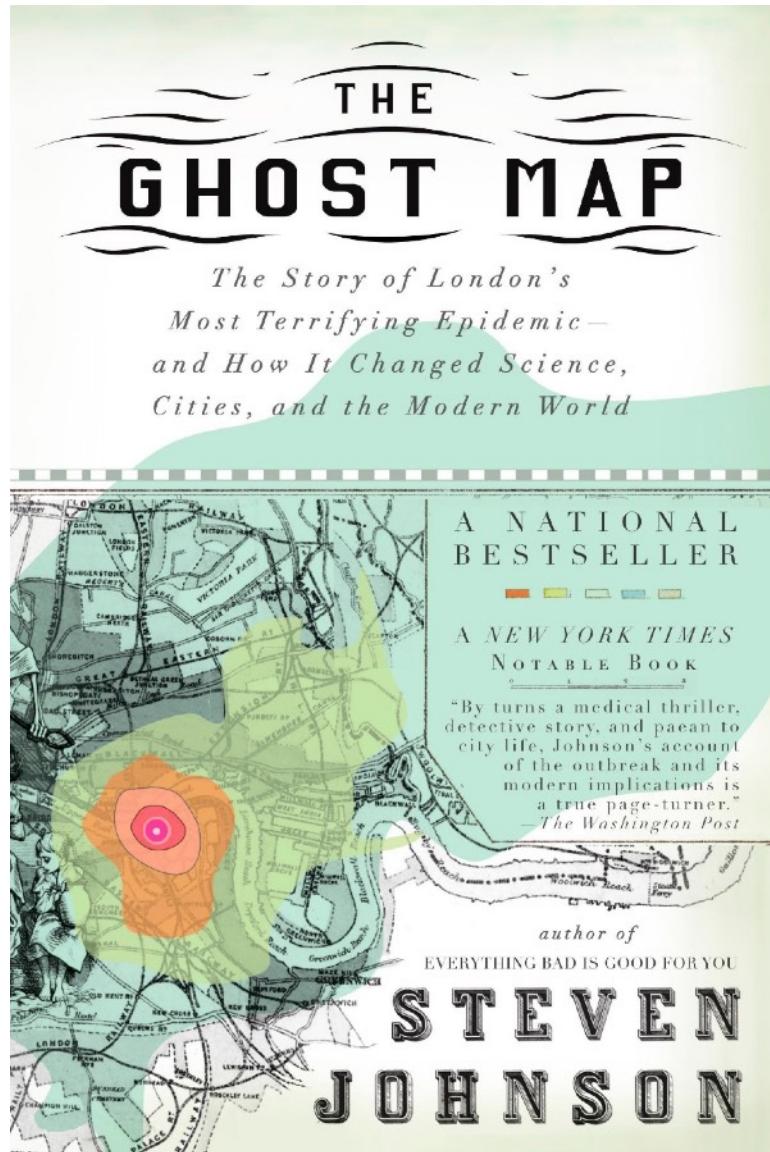
John Snow's map of cholera cases linked to the outbreak



THE GHOST MAP

*The Story of London's
Most Terrifying Epidemic—
and How It Changed Science,
Cities, and the Modern World*





The dominant theory of disease at the time was **miasma** – disease was spread through bad air.

John Snow spent years accumulating evidence which contradicted this theory.

The Broad Street outbreak of 1854 offered the data to support his speculative model of how cholera spread: water.

By itself, data does not tell us anything about the **process by which it was produced**.

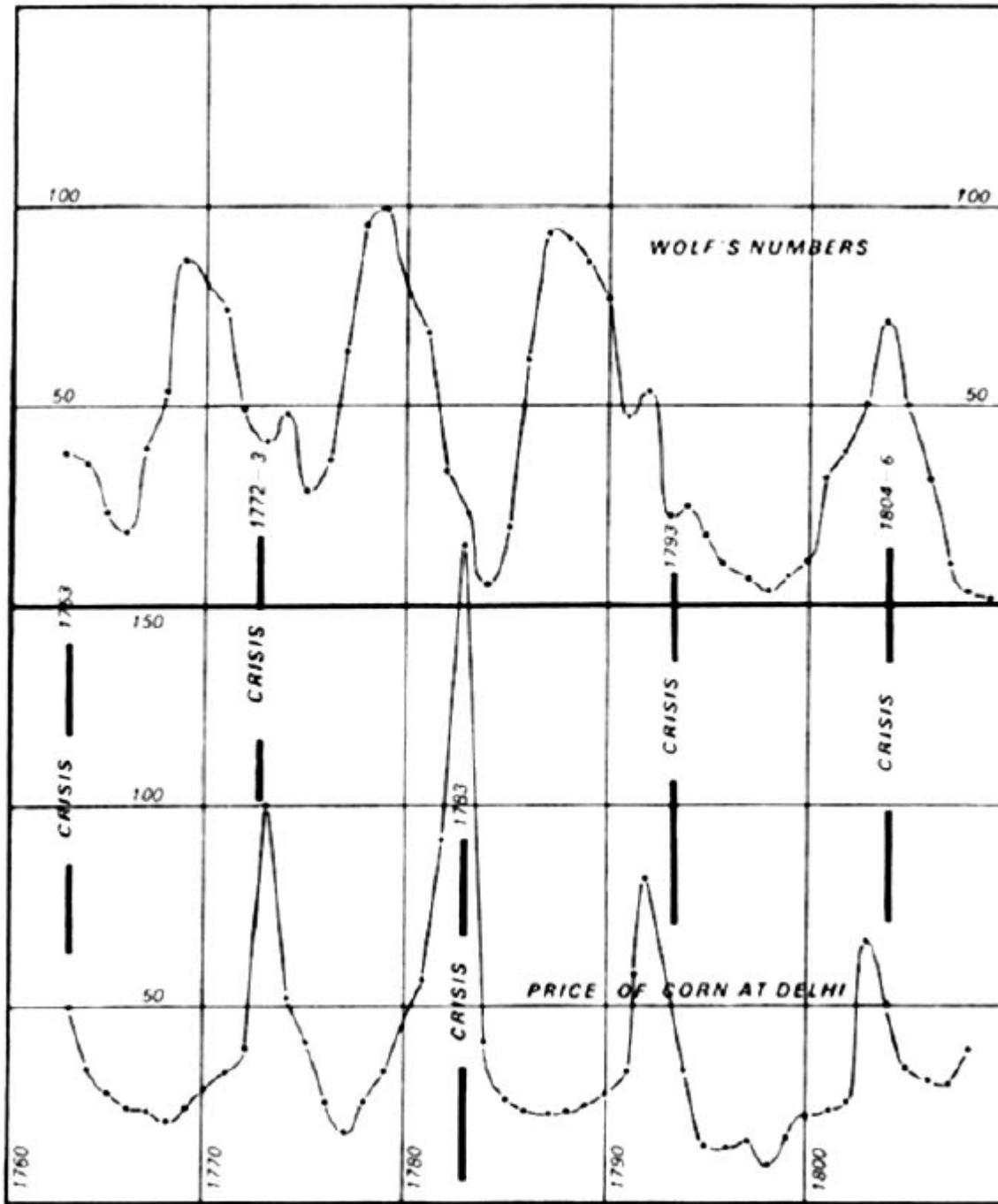
It is incredibly easy to observe data and reach a conclusion about the process that is **completely wrong**.

Being wrong is okay – it is essential to the model building process!

In order to learn from data, we must put our speculations to the test and be willing to be wrong.

It is by figuring out when our models are wrong that we learn.

William Jevons, an economist, was adamant that **sunspots caused economic cycles on earth.**

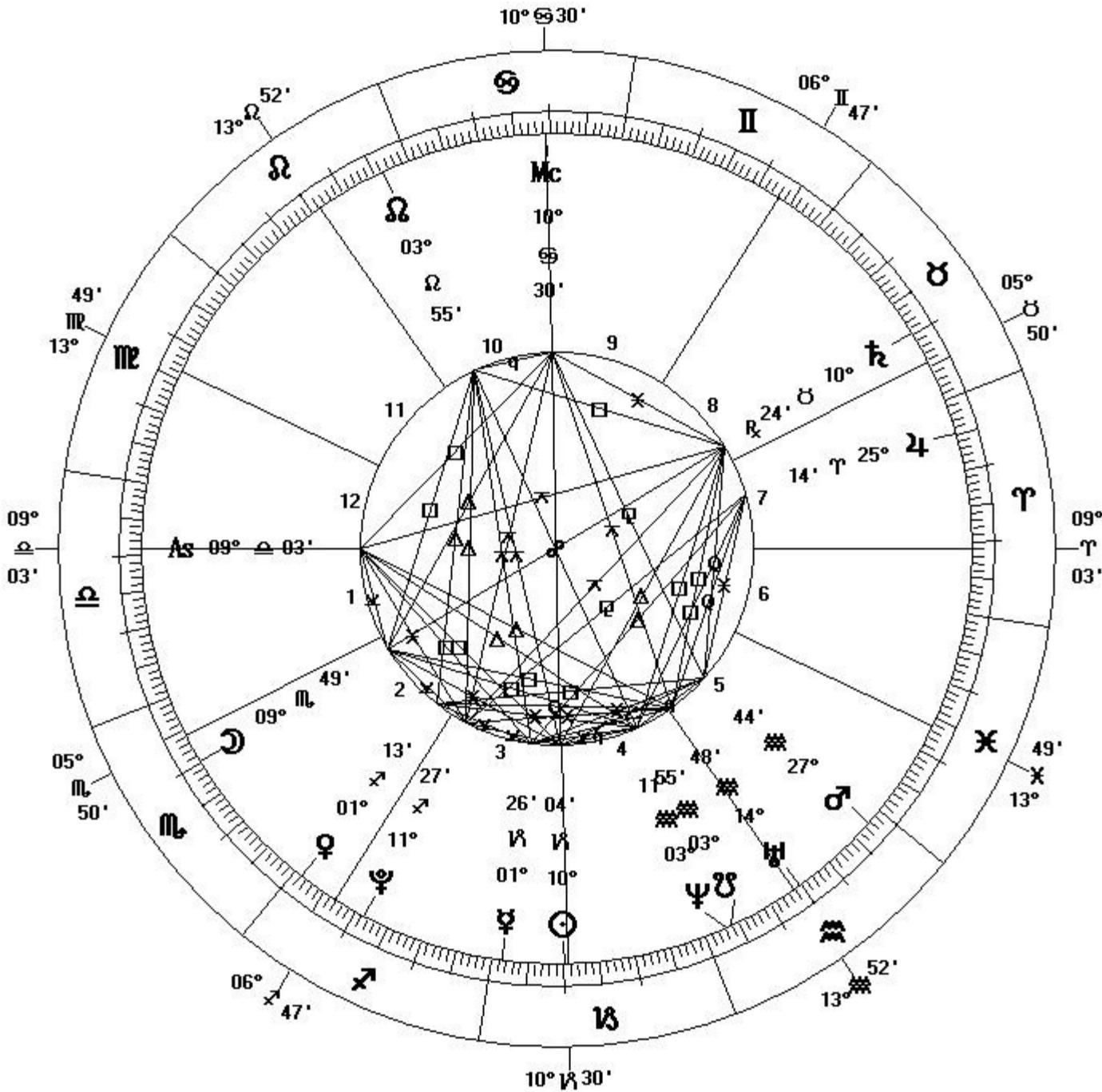


Discovering that our predictions are wrong can be good...

...unless we ignore the data we don't like and **select only the data that shows we're right.**

Astrology was at one point a respected scholarly inquiry: do patterns in the stars determine behavior on earth?

In spite of the lack of evidence, astrologers refused to say no (but they did collect a lot of good data).



Ronald Fisher, the father of statistical science and experimental methods, was adamant that smoking had no effect on rates of cancer.

LETTERS TO THE EDITORS

The Editors do not hold themselves responsible for opinions expressed by their correspondents. No notice is taken of anonymous communications.

Cancer and Smoking

THE curious associations with lung cancer found in relation to smoking habits do not, in the minds of some of us, lend themselves easily to the simple conclusion that the products of combustion reaching the surface of the bronchus induce, though after a long interval, the development of a cancer. If, for example, it were possible to infer that smoking cigarettes is a cause of this disease, it would equally be possible to infer on exactly similar grounds that inhaling cigarette smoke was a practice of considerable prophylactic value in preventing the disease, for the practice of inhaling is rarer among patients with cancer of the lung than with others.

Such results suggest that an error has been made, of an old kind, in arguing from correlation to causation, and that the possibility should be explored that the different smoking classes, non-smokers, cigarette smokers, cigar smokers, pipe smokers, etc., have adopted their habits partly by reason of their personal temperaments and dispositions, and are not lightly to be assumed to be equivalent in their genotypic composition. Such differences in genetic make-up between these classes would naturally be associated with differences of disease incidence without the disease being causally connected with smoking. It would then seem not so paradoxical that the stronger fumes of pipes or cigars should be so much less associated with cancer than those of cigarettes, or that the practice of drawing cigarette smoke in bulk into the lung should have apparently a protective effect.

"If, for example, it were possible to infer that smoking cigarettes is a cause of this disease, it would equally be possible to infer on exactly similar grounds that inhaling cigarette smoke was a practice of considerable prophylactic value in preventing the disease"

Since my letter was written, however, I have received from Dr. Eliot Slater, of the Maudsley Hospital (London, S.E.5), some further data, the greater part of which concern girl twins, and in this way supply a valuable supplement to Verschuer's data, and in which, moreover, a considerable number of pairs were separated at or shortly after birth.

For the resemblance in smoking habit, these female pairs give :

	Alike	Unlike	
Monozygotic	44	9	53
Dizygotic	9	9	18

So far, there is only a clear confirmation of the conclusion from the German data that the monozygotics are much more alike than the dizygotics in their smoking habits. The peculiar value of these data, however, lies in the subdivision of the monozygotic pairs into those separated at birth and those brought up together. These are :

	Alike	Unlike	
Separated	23	4	27
Not separated	21	5	26

Of the nine cases of unlike smoking habit, only four occur among the twenty-seven separated at birth. It would appear that the small proportion unlike among these 53 monozygotic pairs is not to be ascribed to mutual influence.

There is nothing to stop those who greatly desire it from believing that lung cancer is caused by smoking cigarettes. They should also believe that inhaling cigarette smoke is a protection. To believe either is, however, to run the risk of failing to recognize, and therefore failing to prevent, other and more genuine causes.

RONALD A. FISHER

Department of Genetics,
Cambridge.

¹ Fisher, R. A., *Nature*, 182, 108 (1958).

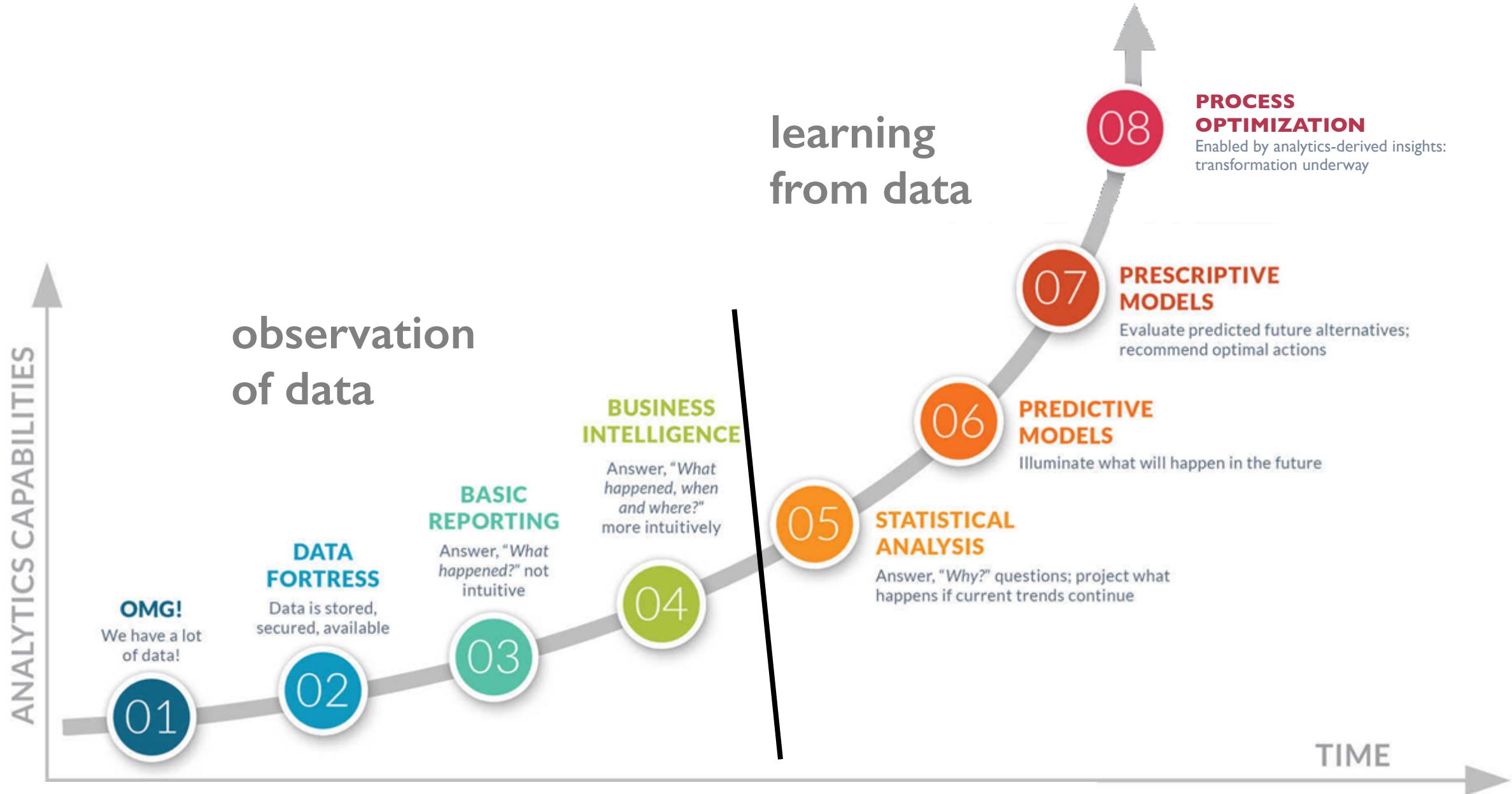
² "Geminus", *New Scientist*, 4, 440 (1958).

Why should we do more than visualization
and reporting? Why build models?

Why should we do more than visualization
and reporting? Why build models?

Our goal is to **learn from data**.
Observation is the start, it is not the end.

Speculating, developing implications, and
putting them to the test is **how we learn**.



learning from data

PROCESS OPTIMIZATION

Enabled by analytics-derived insights:
transformation underway

observation of data

BASIC REPORTING

Answer, "What
happened?" not
intuitive

BUSINESS INTELLIGENCE

Answer, "What
happened, when
and where?"
more intuitively

PREDICTIVE MODELS

Illuminate what will happen in the future

STATISTICAL ANALYSIS

Answer, "Why?" questions; project what
happens if current trends continue

OMG!
We have a lot
of data!

DATA FORTRESS

Data is stored,
secured, available

01

02

03

04

05

06

08

TIME

Predictive models, process optimization -
these all depend on **having learned**
something from data.

Predictive models, process optimization -
these all depend on **having learned
something from data.**

Two things you need for a data science project:

- 1) There is some unknown pattern that would be useful to learn.
- 2) There is data to learn the pattern.

The best way to learn about building
predictive models is to do it.

Let's look at an example in some depth.

The best way to learn about building
predictive models is to do it.

Let's look at an example in some depth.

~~I've got a bad feeling about this~~
This is where the fun begins

2

Building Models of the Universe

^

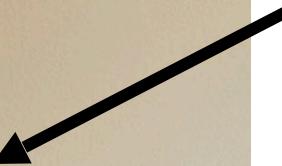
Board Game

Anyone who has worked on a project with me in the last couple years will have noticed two things about my workstation:

Tons of
board
games



Biscuit



**Massive
shelf of
board
games**

Star Wars helmets



Star Wars
boardgames

This massive shelf of board games exists
for two reasons:

- I) I love board games.

This massive shelf of board games exists for two reasons:

- 1) I love board games.
- 2) It is part of an ongoing scientific research agenda: why are some games better than others?

This massive shelf of board games exists for two reasons:

- 1) I love board games.
- 2) It is part of an ongoing scientific research agenda: why are some games better than others?



FLORIDA STATE UNIVERSITY

**AE BUSINESS
SOLUTIONS**

EST. 1949

Funding Provided By

When a global pandemic hits and you're physically separated from people and can't play games for over a year, you do what anyone would do:

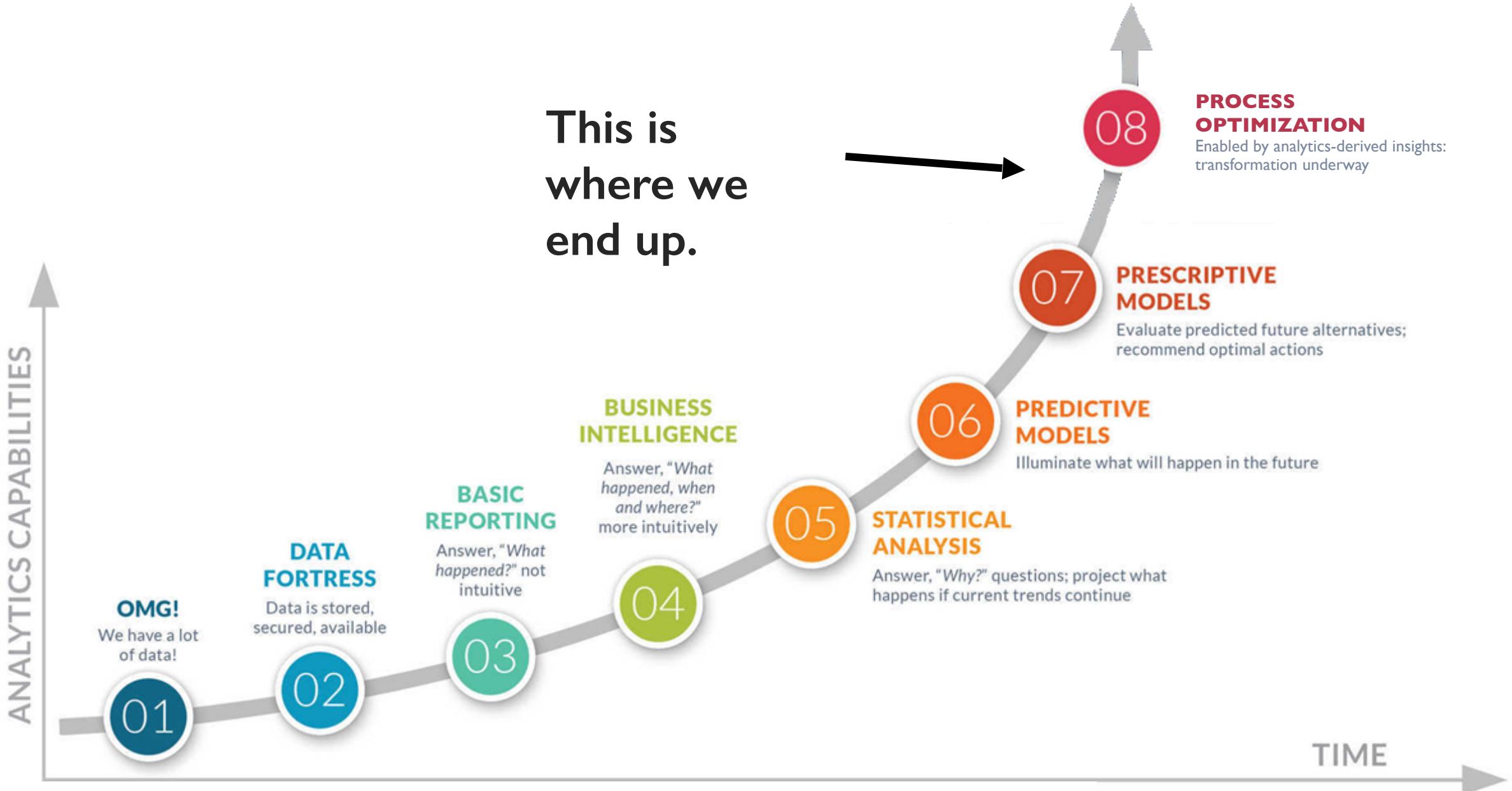
When a global pandemic hits and you're physically separated from people and can't play games for over a year, you do what anyone would do:

Collect data on every boardgame in existence and use the entirety of your scientific training to learn about boardgames so that you can predict upcoming games and recommend them to people so you have something to look forward to when the pandemic is over.

while
ridiculous, this
can illustrate
the life cycle of
a data science
project

When a global pandemic hits and you're physically separated from people and can't play games for over a year, you do what anyone would do:

Collect data on every boardgame in existence and use the entirety of your scientific training to learn about boardgames so that you can **predict upcoming games** and **recommend them to people** so you have something to look forward to when the pandemic is over.



Predicted Geek Ratings for Upcoming Games

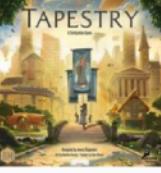
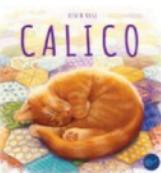
Estimated ratings for boardgamegeek outcomes from predictive models trained prior to 2022

RANK	PUBLISHED	IMAGE	GAME	DESCRIPTION	COMPLEX	RATINGS	AVERAGE	GEEK
1	2024		Gloomhaven: Second Edition (2024)	Gloomhaven: Second Edition is a revised and elevated version of the award-winning core game of Gloomhaven. This is the culmination of everything Isaac Childres and the growing Cephalofair Games team have learned since the initial release of Gloomhaven, including feedback from the community, playtesters, co-designers, and developers. The world, story, and challenging gameplay are all still the ...	3.8	2600	8.8	7.4
2	2024		Galactic Cruise (2024)	Hello, and welcome to Galactic Cruise. Here, we offer our guests something special: the comfort of a luxury cruise with the innovation of space travel. As the first company to offer extended-stay space vacations, we are excited to have you working for us! As a supervisor of this company, you'll be expected to not only build these shuttles and satisfy our guests, but also to help the company th...	4.2	3750	8.1	7.2
3	2024		Santorini: Pantheon Edition (2024)	NOTE: Pantheon Edition includes all material from the Golden Fleece expansion, as well as every promo card and many new cards. A lot of cards are revamped, reworked and rebalanced. Santorini is an accessible strategy game, simple enough for an elementary school classroom while aiming to provide gameplay depth and content for hardcore gamers to explore. The rules are simple. Each turn consists ...	2.0	3350	7.6	6.8
4	2024		Midhalla (2024)	Midhalla is a cooperative dungeoncrawler that focuses on strategic hero mechanics, tower defense elements, and a character-centered story. The game follows the story of Harkon, a newly trained Viking warrior as he faces the invading demon hordes. At the core of Midhalla's gameplay is a strategic hero combat mechanic that eliminates randomness and focuses on complex strategy. Players invest reso...	3.5	1450	8.5	6.8
5	2024		Zombicide: White Death (2024)	Winter is upon us and its dark and cruel night brings forth dreadful menaces, worse than frostbite... Zombicide Medieval Fantasy is back! With new survivors, new zombies, new setting, and new rules! Scale the great walls to your advantage Use the added height of the fortress walls to escape the Zombies' grasp, dump cauldrons of hot tar onto enemies below, and gain superior line of sight! Ar...	3.0	1800	8.0	6.7
6	2024		1848: The Springtime of Nations (2024)	"Do you not feel, through a kind of instinctive intuition that cannot be analyzed but that is certain, that the ground is shaking once again in Europe? Do you not feel—how shall I call it?—that a revolutionary wind is blowing?" – Alexis de Tocqueville, Souvenir (January, 1848) 33 years after the Congress of Vienna, Continental Europe still lives under the reactionary system designed by the Aus...	3.6	1400	8.3	6.6
7	2024		Dragon Eclipse (2024)	Enter a magical world of mythical beasts and ancient dragons in Dragon Eclipse, the exciting new board game for 1-2 players! Uncover the secrets of the eclipse that changed everything and discover the dangerous and treacherous world that awaits you. Explore a vivid and immersive world filled with majestic and unique creatures you can battle, tame, catch, and train. Each creature has a unique d...	2.8	1100	8.5	6.6
8	2024		A Fading Star: Insurgency and Piracy in Somalia (2024)	A Fading Star: Insurgency and Piracy in Somalia is Volume XV of the COIN Series originally designed by Volko Ruhnke. The year is 2007. We reach the climax of the several decades of civil war that have ravaged Somalia since the fall of the dictator Siad Barre in 1991. An African Union coalition, assisted by the Ethiopian army, enters Mogadishu to support a nascent but corrupt Transitional Federal Go...	3.8	1100	8.4	6.5

Predicting which new games are going to be highly rated/popular

Top (Older) Games for rahdo

Rankings based on predictive model trained on user's collection using games released through 2021

RANK	IMAGE	GAME	DESCRIPTION	PR(OWN)	OWN
1		The Magnificent (2019)	The Magnificent is a tightly designed Eurogame from the creators of Santa Maria set in a mystical world beautifully illustrated by French artist Martin Mottet. In the game, players are competing to attract the largest audiences to their shows, featuring magnificent performers. In the process, you must expand your camp by placing Tetris-style tiles on your player board, gather elements needed f...	0.999	yes
2		Tapestry (2019)	Tapestry is a two-hour game for 1-5 players designed by Jamey Stegmaier. Create the civilization with the most storied history, starting at the beginning of humankind and reaching into the future. The paths you choose will vary greatly from real-world events or people – your civilization is unique! In Tapestry, you start from nothing and advance on any of the four advancement tracks (science,...	0.997	yes
3		Among the Stars (2012)	Among the Stars takes place in a war-ravaged galaxy where the warring alien races have declared peace in the wake of a threat with the potential to destroy them all. An Alliance is established to build space stations throughout the galaxy in order to promote trade among the races, strengthen diplomatic relations, and defend against this impending threat. Each player takes the role of one of tho...	0.996	no
4		The City (2011)	In The City, players try to quickly and skillfully build their own city over the rounds, bringing in revenue and victory points (VPs). The cards are both buildings and currency, and the first player to collect 50 victory points wins the game. Victory points are cumulative from round to round, so a fast start yields a big advantage. Each turn, each player places a card from his hand face down, ...	0.994	no
5		Calico (2020)	Calico is a puzzly tile-laying game of quilts and cats. In Calico, players compete to sew the coziest quilt as they collect and place patches of different colors and patterns. Each quilt has a particular pattern that must be followed, and players are also trying to create color and pattern combinations that are not only aesthetically pleasing, but also able to attract the cuddliest cats! Turn...	0.993	yes
6		Walking in Burano (2018)	Burano is a beautiful island of Venice in Italy, known for its colorful houses. Walking through the island, you will see these vibrant houses on both sides of the canal, as well as the personal décor placed by the inhabitants of these houses and shops. Tourists always linger on the street to appreciate and enjoy the view. The colorful houses of Burano need a fresh coat of paint in order to stay...	0.990	yes

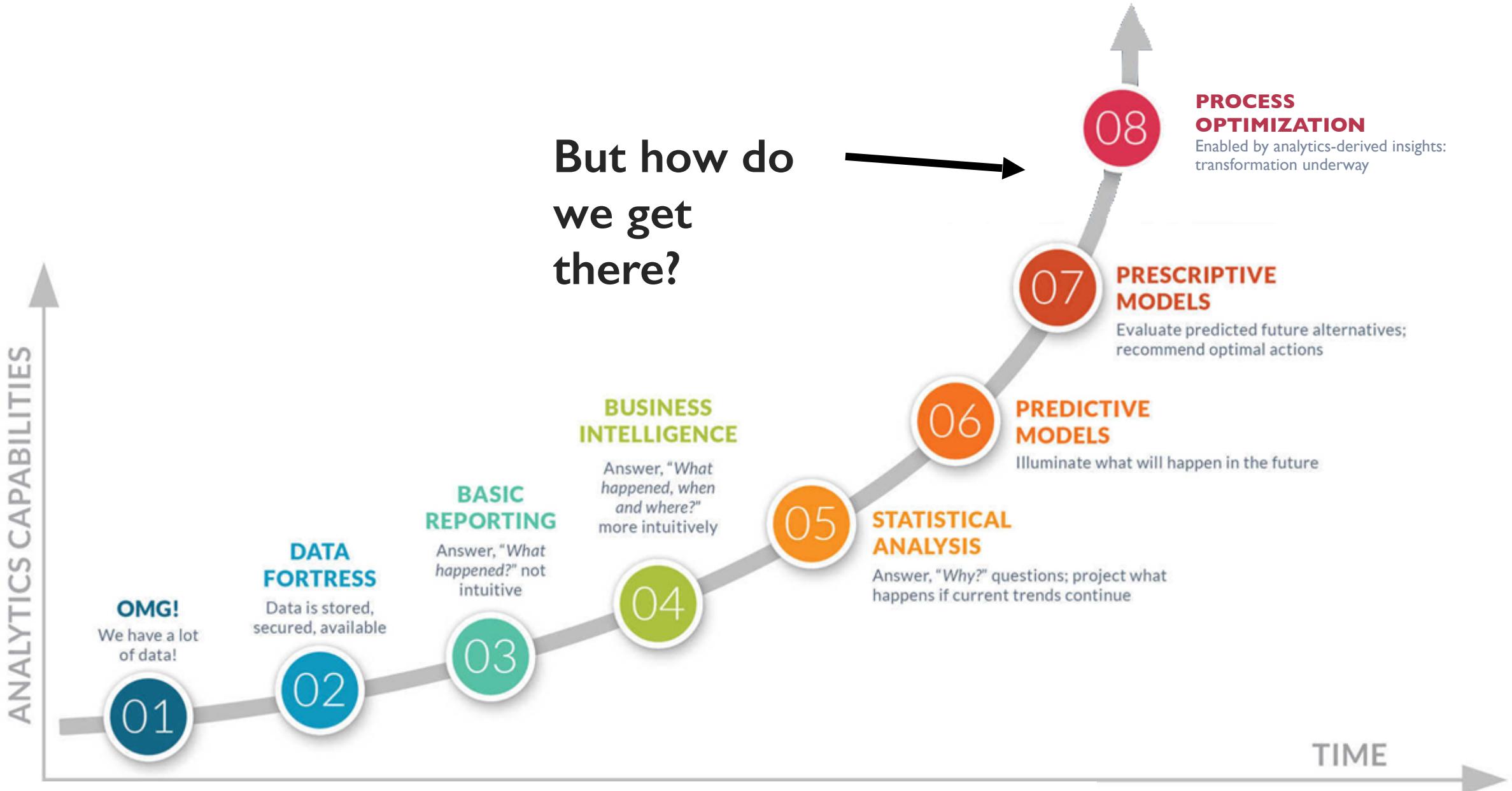
Recommending
games for
individuals
based on their
collections

Top 15 (Newer) Games for GOBBluth89

Rankings based on predictive model trained on user's collection using games released through 2021

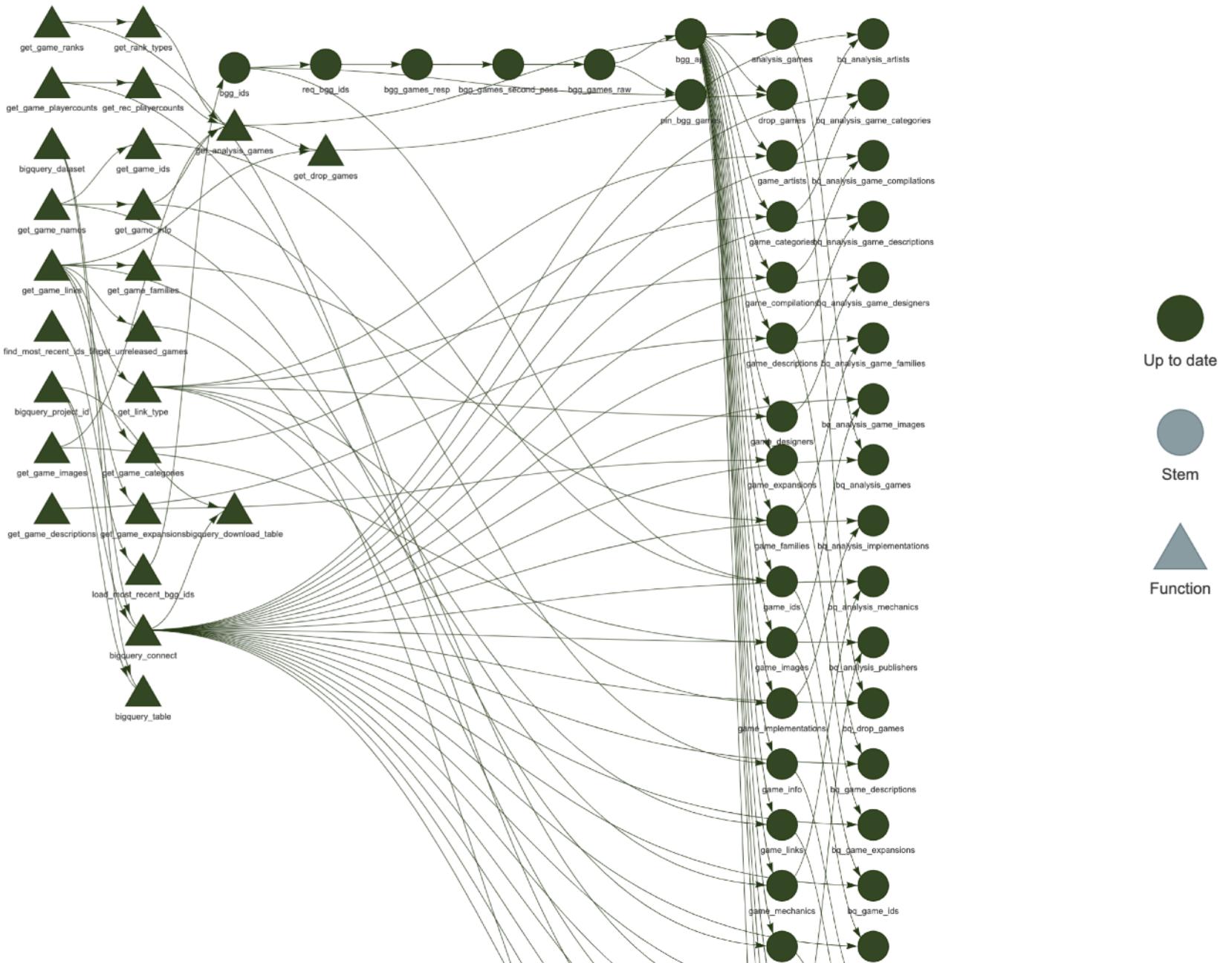
RANK	IMAGE	GAME	DESCRIPTION	PR(OWN)	OWN
1		Unmatched: Jurassic Park - Dr. Sattler vs. T. Rex (2022)	In battle, there are no equals. "Dinosaurs eat man... Woman inherits the earth." The greatest predator the world has ever known is closing in on the tenacious Dr. Sattler. Who has the slightest idea what to expect? In Unmatched: Jurassic Park - Dr. Sattler vs. T. Rex, the massive T rex unleashes fearsome attacks and seems unstoppable while Dr. Sattler makes full use of her surroundings and the a...	0.946	yes
2		The Lord of the Rings: The Card Game - Revised Core Set (2022)	Sometimes, in order to truly appreciate a tale, one must first go back to its beginning. Grand adventures and strong fellowships are important and wonderful, but the first step of any journey is just as important as the last. With that in mind, it's time to return to the beginning of one of the most epic adventures of all... With increased contents and some quality-of-life improvements, this new...	0.827	no
3		Unmatched: Brains and Brawn (2023)	Unmatched: Brains and Brawn, the fifth and final Unmatched Marvel set, features some of Marvel's hottest heroes: Spider-Man, Dr. Strange, and She-Hulk. Spidey swings around the battlefield, using his spider-sense to keep him safe. Dr. Strange has, well, the best card names in the game: Behold the Seven Suns of Cinnibus! And She-Hulk won't think twice about throwing the book — or whatever heavy ...	0.816	no
4		Star Wars: Jabba's Palace - A Love Letter Game (2022)	The lair of Jabba, the Hutt is a place of danger and deception. Players compete using members of the Rebel Alliance and denizens of Jabba's Palace in a game of rebel bravery and vile deceit. Based on the beloved card game and set in the Star Wars galaxy, Star Wars: Jabba's Palace - A Love Letter Game retains the simple and fun mechanics of the original as you utilize the talents of iconic char...	0.780	no
5		The 13th Street Crew (2023)	The 13th Street Crew is a semi-cooperative social deduction game of criminal strategy. The players are low-ranking members of a large criminal organization headed by the Old Don that for all intents and purposes runs this fair city. The players represent fellow crew members occupying the lowest rung in the organization. Most of the players are ambitious and eager to prove they deserve to adv...	0.606	no
6		The Witcher: Old World (2023)	In The Witcher: Old World, you become a witcher — a professional monster slayer — and immerse yourself in the legendary universe of The Witcher franchise. Set years before the saga of Geralt of Rivia, The Witcher: Old World explores a time when monsters roamed the Continent in greater numbers, creating a constant peril that required the attention of expertly trained monster slayers, known as w...	0.550	no

Recommending
games for
individuals
based on their
collections



**But how do
we get
there?**

Data Integration



Data Warehousing

Explorer + ADD

Type to search

Viewing workspace resources.
SHOW STARRED ONLY

- gcp-analytics-326219
 - Saved queries (5)
 - External connections
- bgg
 - active_bgg_rankings...
 - active_game_ids
 - active_games_daily
 - active_games_info
 - analysis_drop_games
 - analysis_game_artists
 - analysis_game_cate...
 - analysis_game_com...
 - analysis_game_desc...
 - analysis_game_desi...
 - analysis_game_famil...
 - analysis_game_imag...
 - analysis_game_impl...
 - analysis_game_mec...
 - analysis_game_publi...
 - analysis_games
 - analysis_unreleased...
 - api_game_descriptio...
 - api_game_expansions
 - api_game_ids
 - api_game_images
 - api_game_info
 - api_game_links
 - api_game_names
 - api_game_playercou...
 - api_game_ranks

Untitled 2

```
1 SELECT * FROM `gcp-analytics-326219.bgg.analysis_game_families` LIMIT 1000
```

Query results

JOB INFORMATION		RESULTS		JSON	EXECUTION DETAILS		CHART	PREVIEW	EXECUTION GRAPH	
Row	game_id	load_ts	family_type		id	type			family_value	
2	35503	2023-01-07 15:43:08 UTC	Game		3867	family			Ra	
3	12589	2023-01-07 15:43:08 UTC	Game		3867	family			Ra	
4	59753	2023-01-07 15:43:08 UTC	Game		3867	family			Ra	
5	12	2023-01-17 11:27:49 UTC	Game		3867	family			Ra	
6	35503	2023-01-17 11:27:49 UTC	Game		3867	family			Ra	
7	12589	2023-01-17 11:27:49 UTC	Game		3867	family			Ra	
8	59753	2023-01-17 11:27:49 UTC	Game		3867	family			Ra	
9	299	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
10	5369	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
11	8495	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
12	12715	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
13	28096	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
14	31167	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
15	31583	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
16	269227	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
17	284578	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
18	327071	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
19	339932	2023-01-07 15:43:08 UTC	Game		6324	family			DBx	
20	299	2023-01-17 11:27:49 UTC	Game		6324	family			DBx	
21	5369	2023-01-17 11:27:49 UTC	Game		6324	family			DBx	

Data Visualization & Reporting

Find a Boardgame!

About:
This dashboard is for filtering and examining boardgames using data from boardgamegeek.com.

Select from the filters below to find games that might be a good fit for you.

Recommended Player Count
1, 2, 3, 4, 5, 6, 7, 8, 8+

Select Complexity
Low (1-2), Moderate (2-3), H

Select Playing Time
<30 min, 30-60 min, 60-120+

Select a Category
Any Category

Select Cooperative
Competitive, Cooperative, S

Data last refreshed on 2022-05-23

Table of Games

Show 20 entries Search:

Published	Name	Recommended Players	Complexity	Geek Rating	Community Rating
2017	Gloomhaven	1,2,3	3.88	8.49	8.72
2015	Pandemic Legacy: Season 1	2,3,4	2.83	8.43	8.58
2018	Brass: Birmingham	2,3,4	3.9	8.43	8.66
2016	Terraforming Mars	1,2,3,4	3.25	8.26	8.41
2020	Gloomhaven: Jaws of the Lion	1,2,3	3.61	8.26	8.63
2017	Twilight Imperium: Fourth Edition	3,4,5,6	4.26	8.25	8.66
2017	Gaia Project	1,2,3,4	4.37	8.17	8.46
2016	Star Wars: Rebellion	2	3.73	8.17	8.42
2017	Spirit Island	1,2,3	4.04	8.15	8.37
2011	War of the Ring: Second Edition	2	4.19	8.14	8.52
2015	Through the Ages: A New Story of Civilization	2,3	4.42	8.13	8.36
2016	Great Western Trail	2,3	3.71	8.12	8.28
2005	Twilight Struggle	2	3.6	8.11	8.27
2021	Ark Nova	1,2,3	3.71	8.09	8.69
2016	Scythe	1,2,3,4,5	3.43	8.05	8.21
2020	Dune: Imperium	1,2,3,4	3	8.04	8.33
2011	The Castles of Burgundy	2,3	3	8	8.12
2015	7 Wonders Duel	2	2.22	7.98	8.11
2007	Brass: Lancashire	2,3,4	3.86	7.97	8.19
2013	Concordia	2,3,4,5	3.01	7.97	8.12

Showing 1 to 20 of 19,194 entries

Previous 1 2 3 4 5 ... 960

Complexity and Community Rating

Community Rating

Complexity

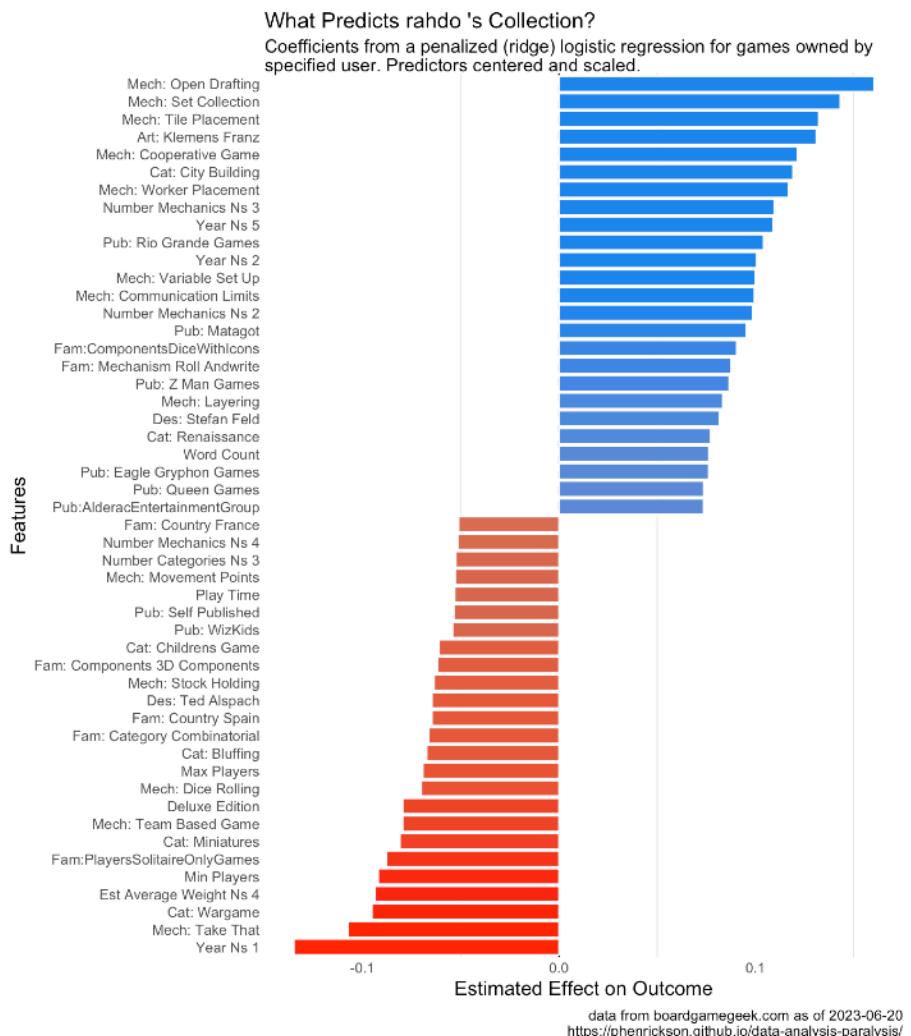
Number of Games by Category

Category	Count
Card Game	~3500
Wargame	~300
Fantasy	~200
Dice	~100
Party Game	~100
Science Fiction	~100
Fighting	~100
Abstract Strategy	~100
Economic	~100
Children's Game	~100
Animals	~100
World War II	~100
Adventure	~100
Bluffing	~100
Humor	~100
Deduction	~100
Miniatures	~100
Medieval	~100
Action / Diversity	~100
Exploration	~100
Movies / TV / Radio theme	~100
Ancient	~100
Real-time	~100
Horror	~100
Puzzle	~100

Category

Number of Games

Predictive Models



1.3 Designer Partial Effects

Which designers have the largest effect on a game's average? User ratings?

We could simply look at the games published by each designer and then take their average rating/average number of user ratings. However, this doesn't account for the fact that some designers design more complex games than others, or that recently published games tend to have higher averages.

To get an estimate of each designer's partial effect on an outcome, I run penalized regressions (lasso) of the BGG average/users rated on dummies for individual game designers along with the game's weight and effects for year published.

I handle time effects in two different ways, first creating an indicator for games published before 1900, then fitting cubic splines to the truncated publishing range of years 1900 to present.

$$\text{Average} = \beta_0 + \text{Weight} * \beta_1 + \text{Designer}_1 * \beta_2 + \text{Designer}_2 * \beta_3 * \text{Designer}_k \dots \text{Designer}_n * \beta_n + \text{YearPublished}$$

Show 15 entries Search:

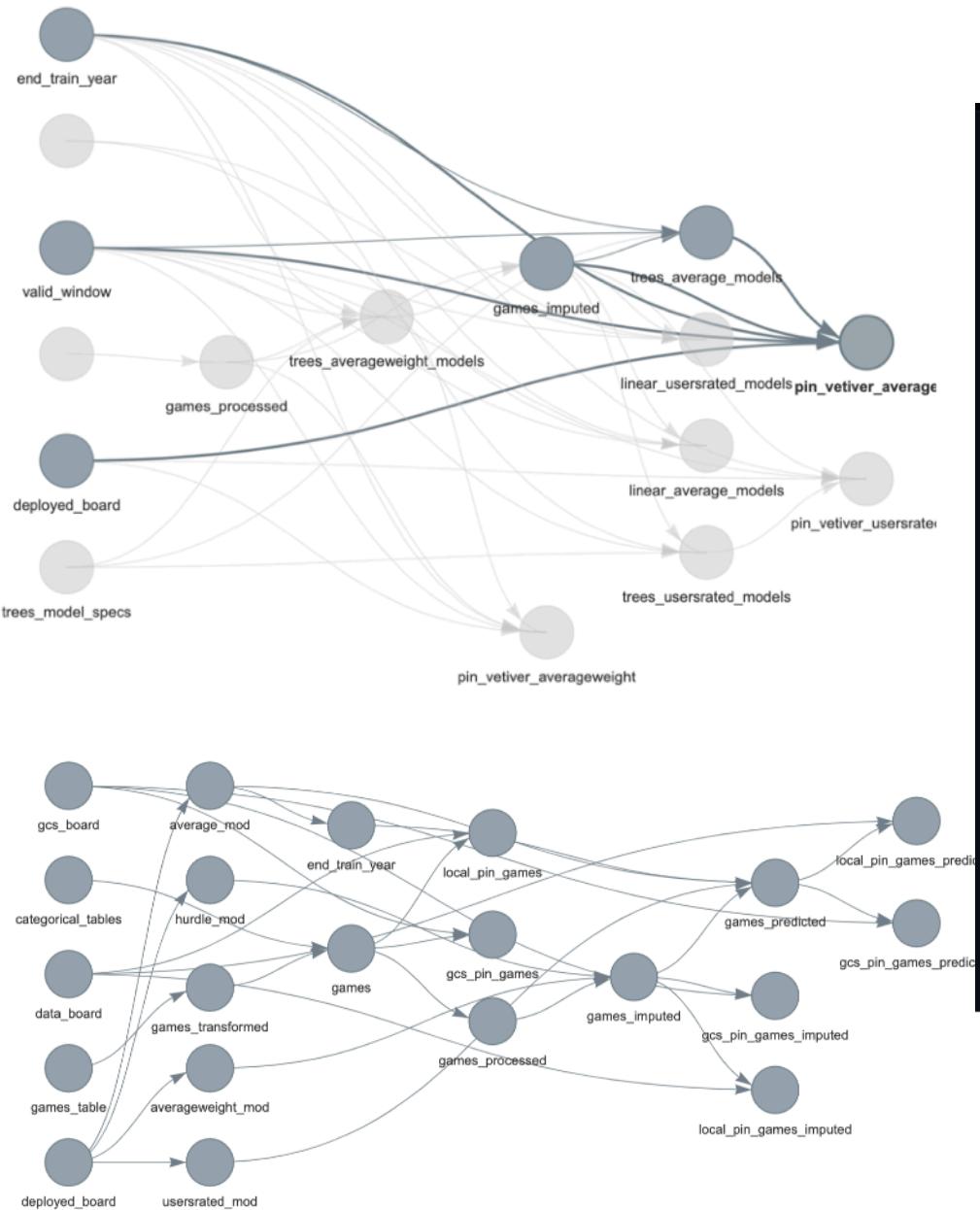
Designer Games Average Ratings

Designer	Games	Average	Ratings
All	10 ... 1494	All	All
Jamey Stegmaier	10	1.254	26.08
Matt Leacock	21	0.86	10.64
Mac Gerdts	12	0.987	8.98
Corey Konieczka	28	0.859	8.47
Vlaada Chvátil	32	0.793	8.4
Vital Lacerda	13	1.092	7.85
Stefan Feld	42	0.744	6.9
Alexander Pfister	24	0.759	6.67
Emerson Matsuuchi	15	0.654	6.42
Wolfgang Warsch	21	0.596	6.28
Philippe Keyaerts	10	0.582	5.35
Dominic Crapuchettes	15	0.505	5.35
Raphaël Guiton	13	0.692	5.28
Vladimír Suchý	12	0.778	5.21
Tim Fowers	10	0.6	5.06

Showing 1 to 15 of 511 entries (filtered from 1,309 total entries)

Previous 2 3 4 5 ... 35 Next

ModelOps & CI/CD



Screenshot of a GitHub repository named `bgg_models` (Public). The repository has 8 branches and 0 tags. The current branch is `dev-update-bgg...`, which is 9 commits ahead and 7 commits behind `main`. The commit history is as follows:

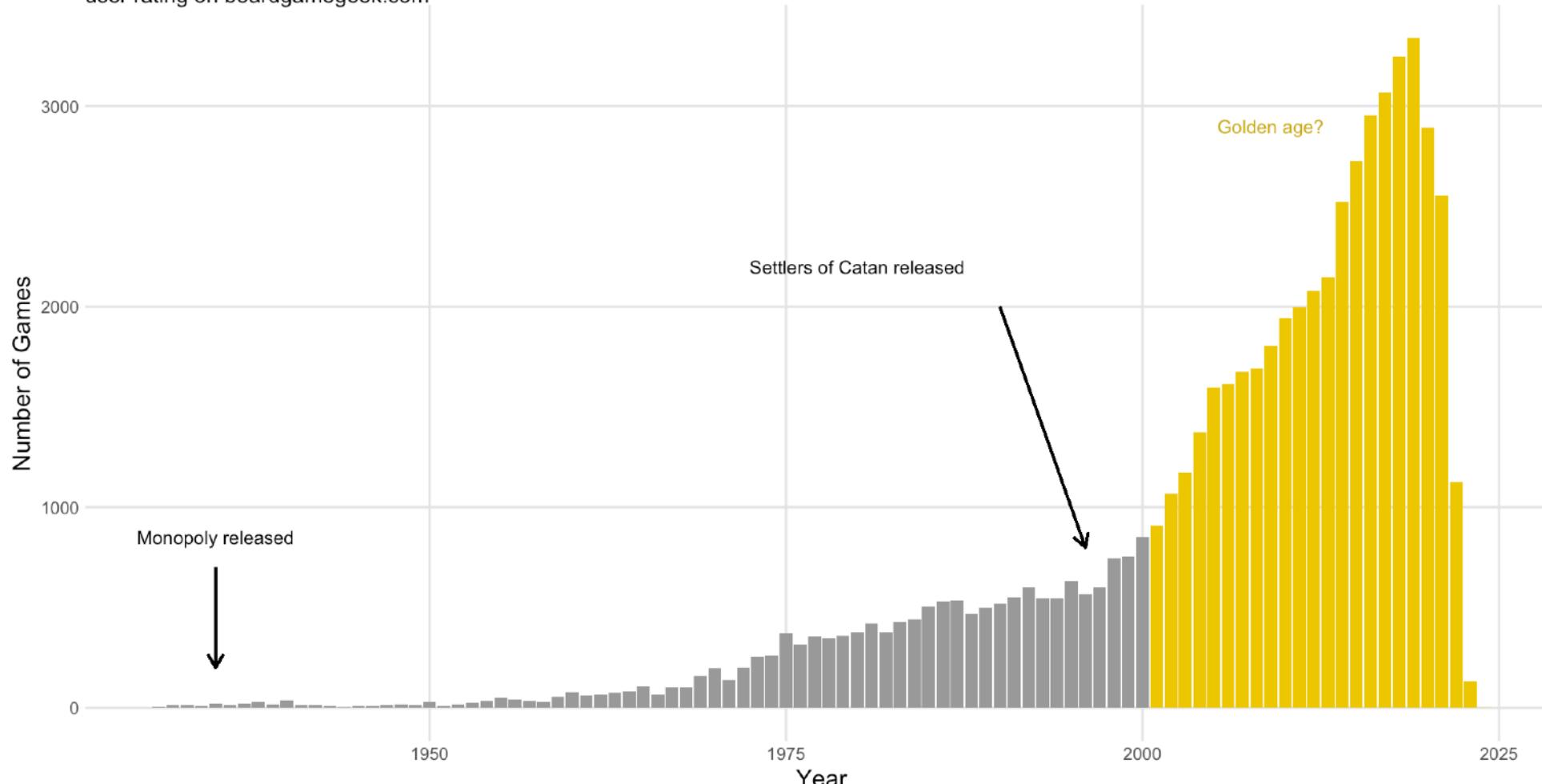
Commit	Message	Date
e82c6a5	phenrickson folded predictions into data pipeline and started reports pipeline	last week
...	folded predictions into data pipeline and started reports pipeline	last week
...	added pipelines with targets to run data, training, and prediction jobs	last week
...	folded predictions into data pipeline and started reports pipeline	last week
...	fixed issue with markdown code inclusion	2 months ago
...	added pipeline to make imputed data for models	2 months ago
...	updated with most recent bgg run	last month
...	folded predictions into data pipeline and started reports pipeline	last week
...	added initial gitignore with standard R files to ignore and addition...	5 months ago
...	updated readme	2 months ago
...	folded predictions into data pipeline and started reports pipeline	last week
...	folded predictions into data pipeline and started reports pipeline	last week
...	folded predictions into data pipeline and started reports pipeline	last week
...	folded predictions into data pipeline and started reports pipeline	last week
...	added pipelines with targets to run data, training, and prediction jobs	last week
b6f3a2d	adding Rproject file	5 months ago
...	added pipelines with targets to run data, training, and prediction jobs	last week

These are peripheral; what matters
for any project is the science.

We start where any project begins:
observation.

A Golden Age of Board Games?

Number of boardgames released by year since 1930, filtering to games with at least one user rating on boardgamegeek.com



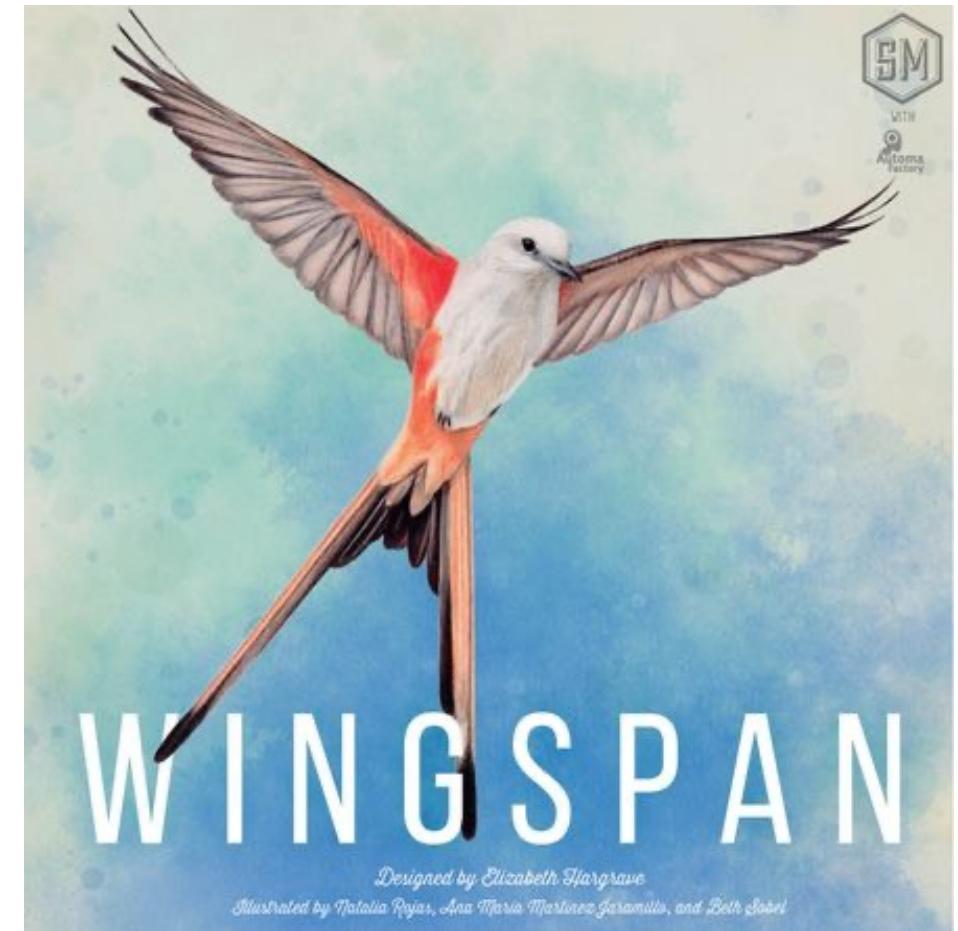
Data from boardgamegeek.com as of 2022-04-17
Analysis at phenrickson.github.io/data-analysis-paralysis/boardgames.html

**Board games exploded in popularity in the mid 2000s;
thousands of new games come out every year.**



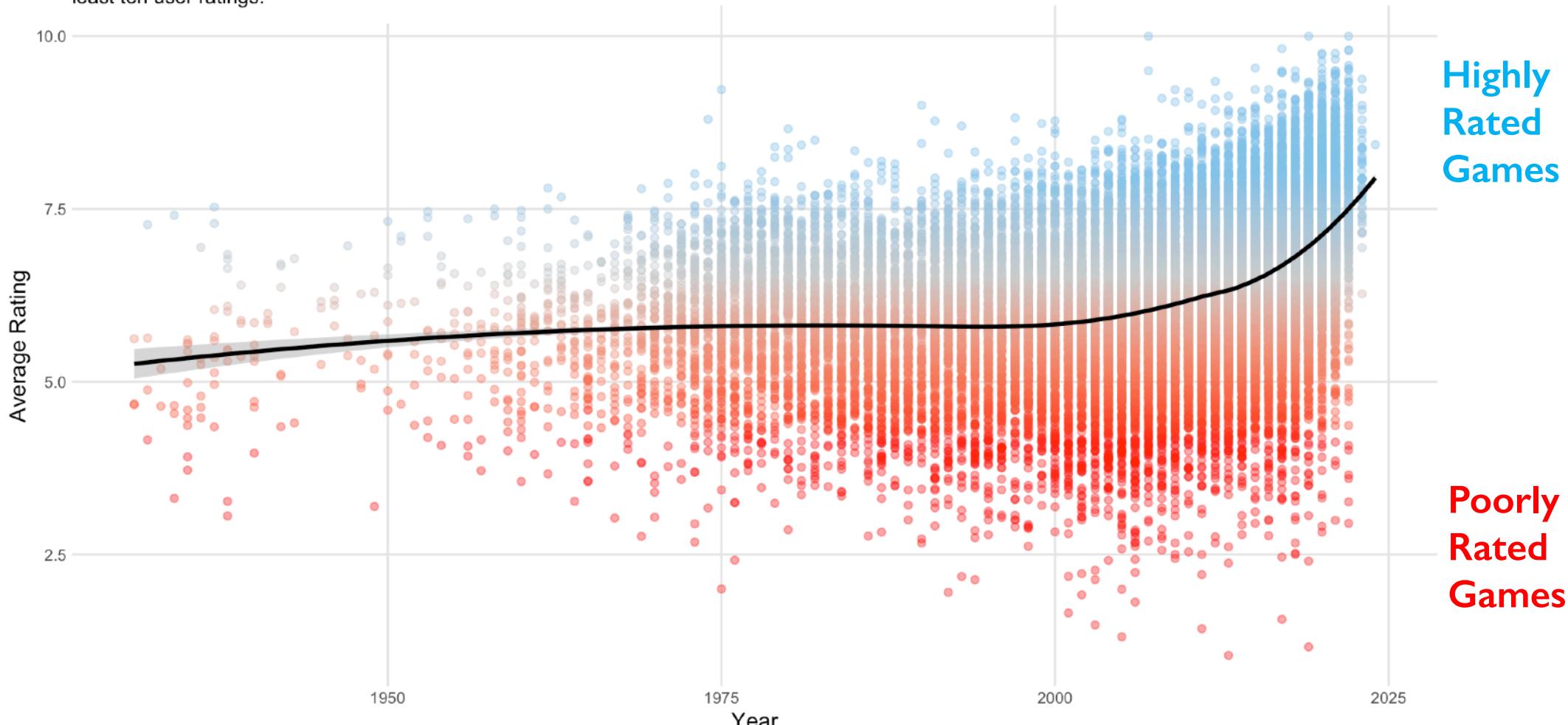
Released in 2007
Over 5 million copies sold
Play as the CDC facing a
novel virus spreading — oh God

Released in 2019
Over 1 million copies sold
Build an aviary and get the
prettiest and best birds



Are Games Getting Better?

Boardgamegeek average rating for games released since 1930, filtering to games with at least ten user ratings.

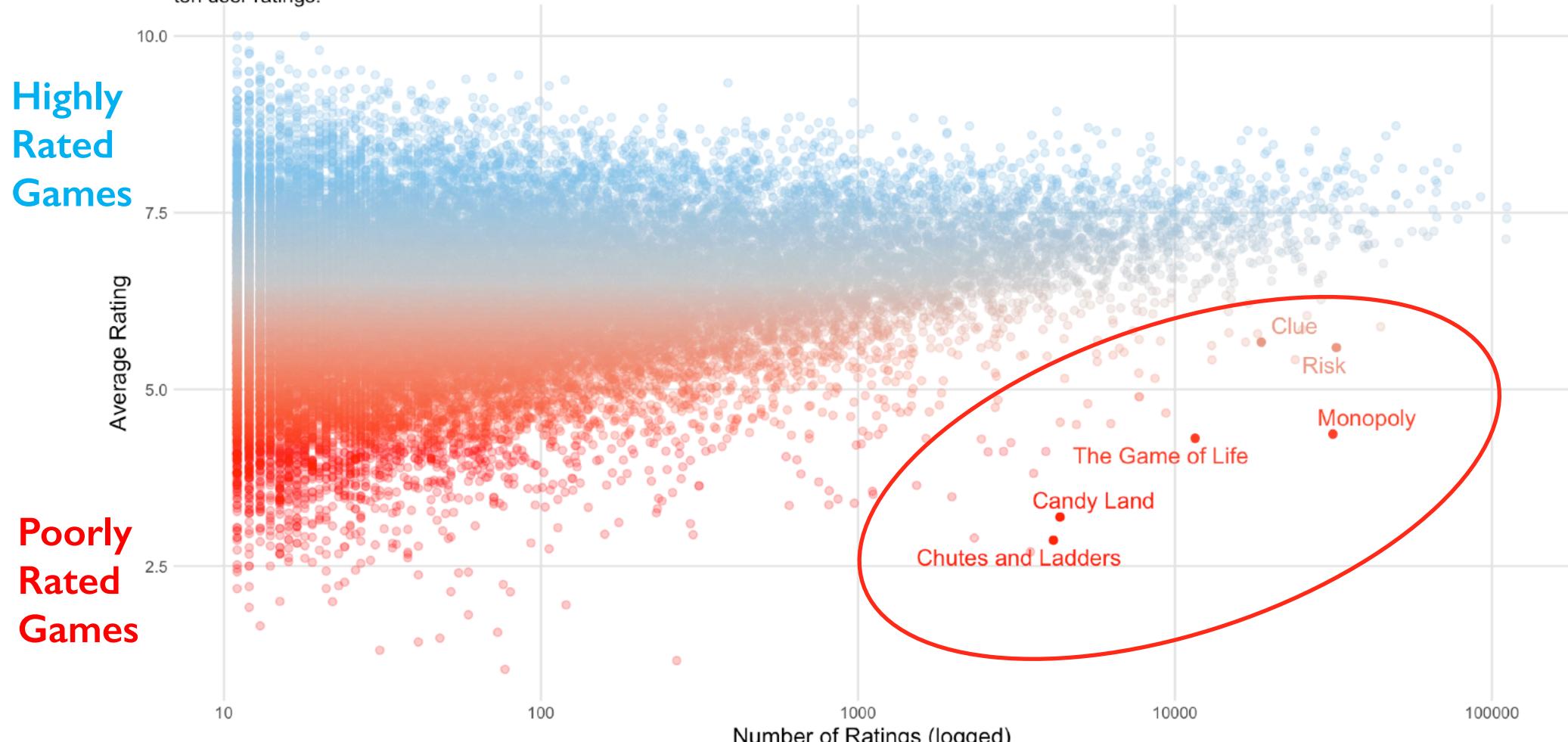


Data from boardgamegeek.com as of 2022-04-17
Analysis at phenrickson.github.io/data-analysis-paralysis/boardgames.html

New games tend to be higher rated than old games –
games themselves might* be getting better.

Many of the Games You Know Are Bad

Boardgamegeek average rating vs number of user ratings, filtering to games with at least ten user ratings.



Data from boardgamegeek.com as of 2022-04-17
Analysis at phenrickson.github.io/data-analysis-paralysis/boardgames.html

Many of the games you know are rated quite poorly by the boardgame community.

If you look at enough board games, you start to ask questions: **why are some games better than others?**

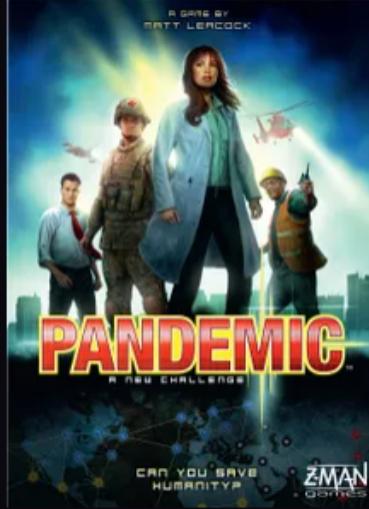
There are so many games out there that it can be overwhelming trying to find games you or your friend might like.



I'M BOARD!
GAMES & FAMILY FUN

How can I help people find games?

I can **collect data** on every game ever published and build some **interactive dashboards** to let people explore the data.



REIMPLEMENTED BY: PANDEMIC LEGACY:... + 10 MORE RANK: OVERALL 119 STRATEGY 125 FAMILY 19

7.6

Pandemic (2008)

Your team of experts must prevent the world from succumbing to a viral pandemic. ↗

112K Ratings & 18K Comments · GeekBuddy Analysis

2–4 Players

Community: 1–4 — Best: 4

45 Min

Playing Time

Age: 8+

Community: 10+

Weight: 2.41 / 5

'Complexity' Rating ↗

Alternate Names: EPIZOotic, Pandemic: 10th Anniversary Edition, 19 + ↗ more

Designer: Matt Leacock

Artist: Josh Cappel, Christian Hanisch, Régis Moulun, Chris Quilliams, Tom Thiel

Publisher: Z-Man Games + 34 more

[See Full Credits](#)

Year Released

2008

Mechanisms

Action Points

Designer

[Matt Leacock](#)

Cooperative Game

Solo Designer

N/A

Hand Management

Artists

[Josh Cappel](#)
[Christian Hanisch](#)
[Régis Moulun](#)
[Chris Quilliams](#)
[Tom Thiel](#)

Point to Point Movement

Set Collection

Trading

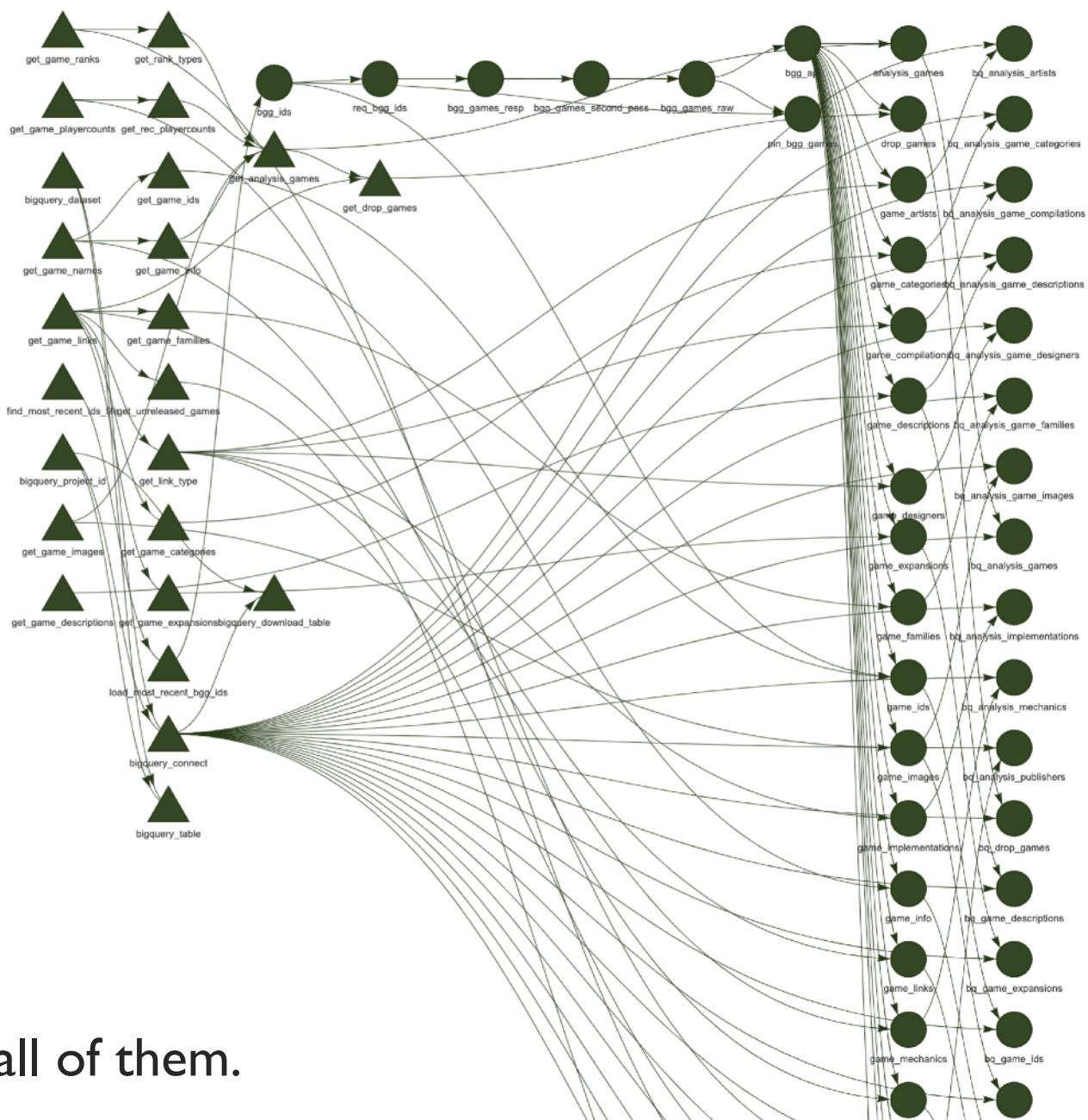
Variable Player Powers

Every single game has a page with data that looks like this.

```

25
26 # functions relating to biggquery
27 tar_source(here::here("src", "data", "connect_to_bigquery.R"))
28
29 # functions to load data
30 tar_source(here::here("src", "data", "load_data.R"))
31
32 list(
33   # scrape universe of bgg ids with python
34   tar_target(
35     name = scraped_bgg_ids,
36     command = reticulate::source_python(here::here("src", "data", "scrap"
37     # only run if its been more than 6 days since previous run
38     cue = tar_cue_age(
39       name = scraped_bgg_ids,
40       age = as.difftime(6, units = "days"))
41   ),
42   # load most recently scraped bgg ids
43   tar_target(
44     name = bgg_ids,
45     command =
46       here("data", "raw") %>
47       load_most_recent_bgg_ids() %>
48       arrange(page)
49   ),
50   # make unique ids for request to api
51   tar_target(
52     name = req_bgg_ids,
53     command =
54       unique(bgg_ids$game_id)
55   ),
56   # submit ids to api in batches
57   tar_target(
58     name = bgg_games_resp,
59     command =
60       bggUtils::get_bgg_games(req_bgg_ids,
61                               batch_size = 500,
62                               tidy = T,
63                               toJSON = F)
64   ),
65   # second pass for any missed ids
66   tar_target(
67     name = bgg_games_second_pass,

```



We'll scrape all of them.

Explorer + ADD

Type to search

Viewing workspace resources.

SHOW STARRED ONLY

- gcp-analytics-326219
 - Saved queries (5)
 - External connections
- bgg
 - active_bgg_rankings...
 - active_game_ids
 - active_games_daily
 - active_games_info
 - analysis_drop_games
 - analysis_game_artists
 - analysis_game_cate...
 - analysis_game_com...
 - analysis_game_desc...
 - analysis_game_desi...
 - analysis_game_famil...
 - analysis_game_imag...
 - analysis_game_Impl...
 - analysis_game_mec...
 - analysis_game_publi...
 - analysis_games
 - analysis_unreleased...
 - api_game_descriptio...
 - api_game_expansions
 - api_game_ids
 - api_game_images
 - api_game_info
 - api_game_links
 - api_game_names
 - api_game_playercou...
 - api_game_ranks

Untitled 2

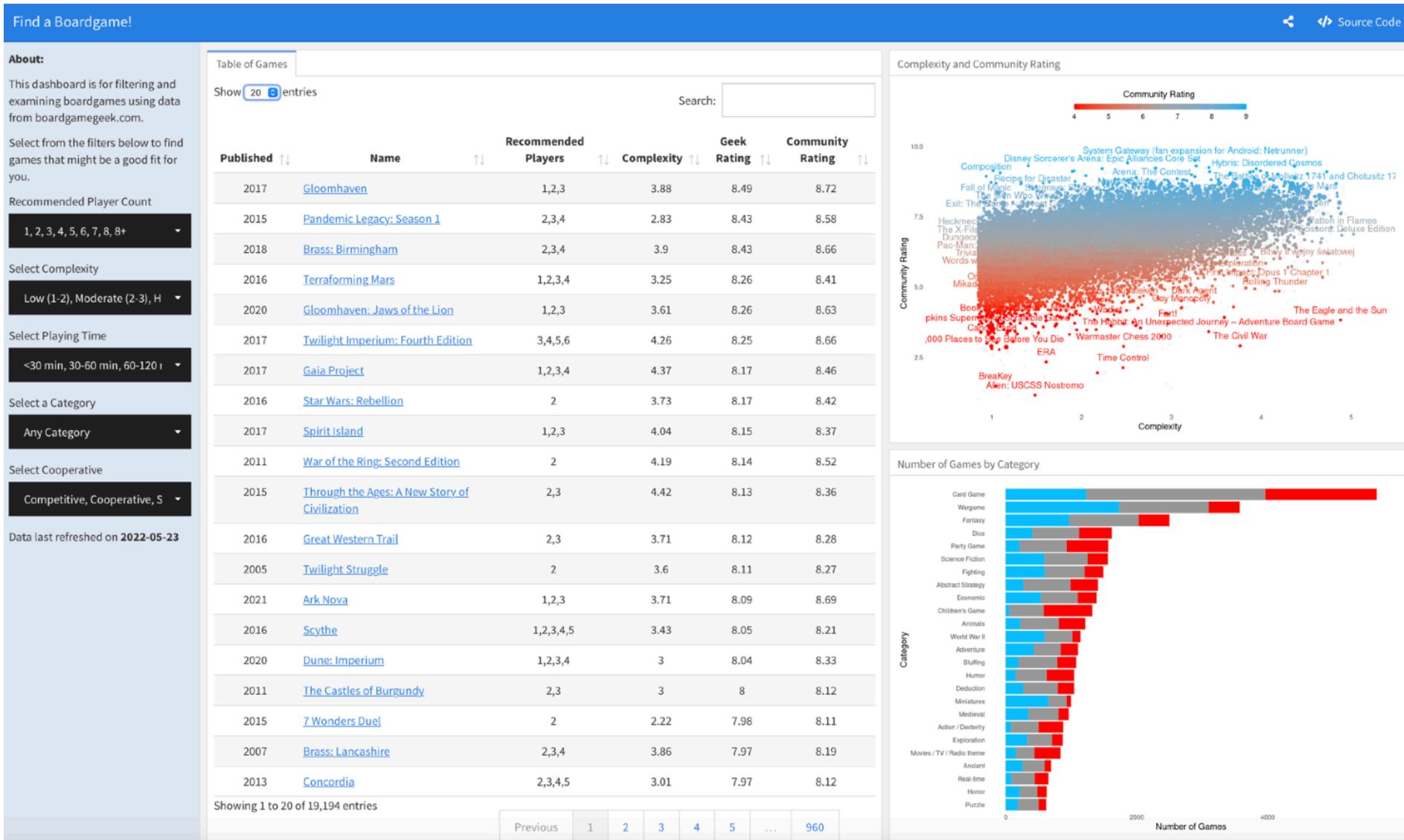
RUN SAVE SHARE SCHEDULE MORE

```
1 SELECT * FROM `gcp-analytics-326219.bgg.analysis_game_families` LIMIT 1000
```

Query results

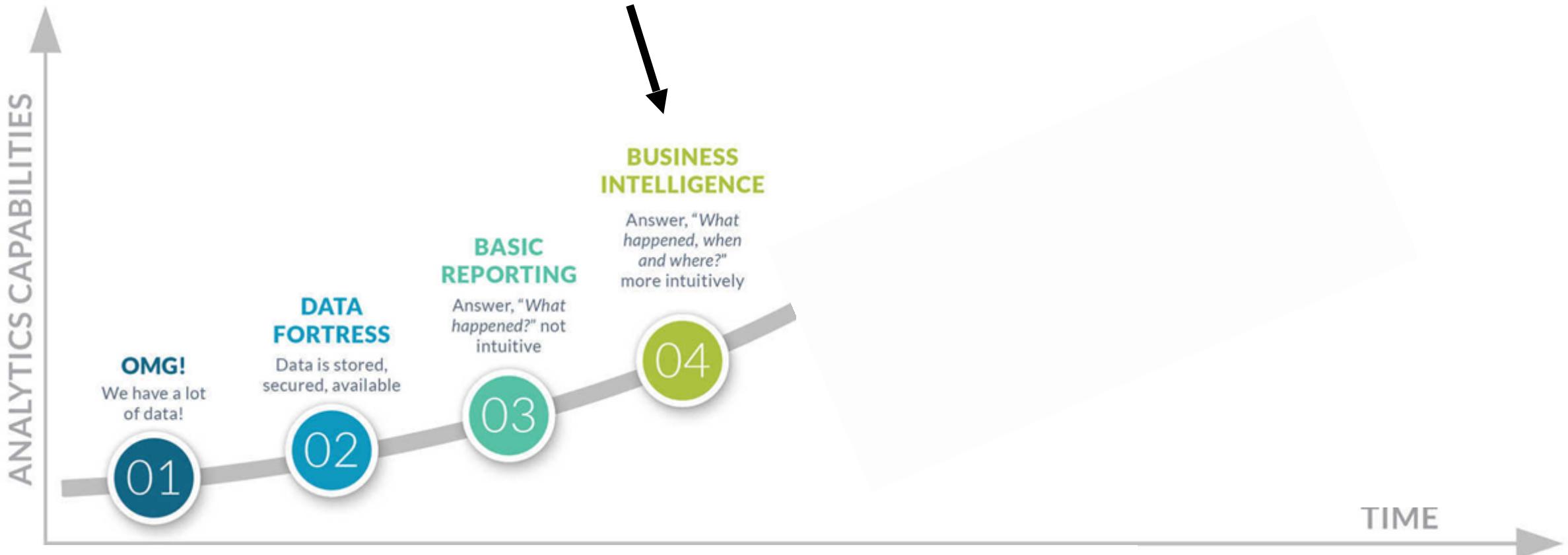
JOB INFORMATION		RESULTS		JSON	EXECUTION DETAILS		CHART	PREVIEW	EXECUTION GRAPH	
Row	game_id	load_ts	family_type		id	type		family_value		
2	35503	2023-01-07 15:43:08 UTC	Game		3867	family		Ra		
3	12589	2023-01-07 15:43:08 UTC	Game		3867	family		Ra		
4	59753	2023-01-07 15:43:08 UTC	Game		3867	family		Ra		
5	12	2023-01-17 11:27:49 UTC	Game		3867	family		Ra		
6	35503	2023-01-17 11:27:49 UTC	Game		3867	family		Ra		
7	12589	2023-01-17 11:27:49 UTC	Game		3867	family		Ra		
8	59753	2023-01-17 11:27:49 UTC	Game		3867	family		Ra		
9	299	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
10	5369	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
11	8495	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
12	12715	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
13	28096	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
14	31167	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
15	31583	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
16	269227	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
17	284578	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
18	327071	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
19	339932	2023-01-07 15:43:08 UTC	Game		6324	family		DBx		
20	299	2023-01-17 11:27:49 UTC	Game		6324	family		DBx		
21	5369	2023-01-17 11:27:49 UTC	Game		6324	family		DBx		

And we'll store everything in a cloud data warehouse
(on the cheap).



Then we can deploy somewhat-ugly-open-source dashboards to let people explore the data.

..alright, so
we're here.



But the data, by itself, can't answer
some of our most important questions:

But the data, by itself, can't answer some of our most important questions:

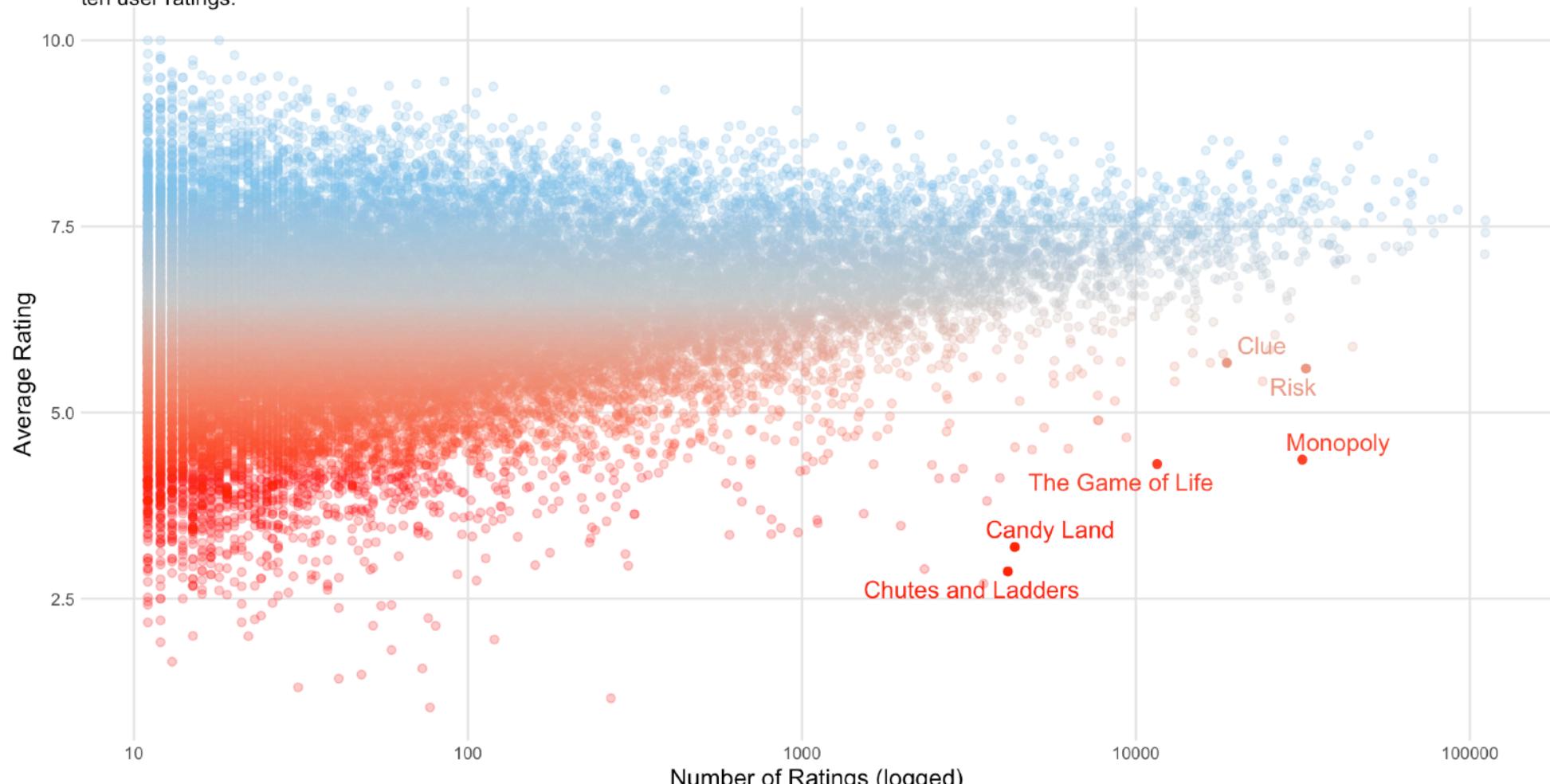
Why are some games highly rated but not others?

What new games are likely to be good?

What games will a specific person like?

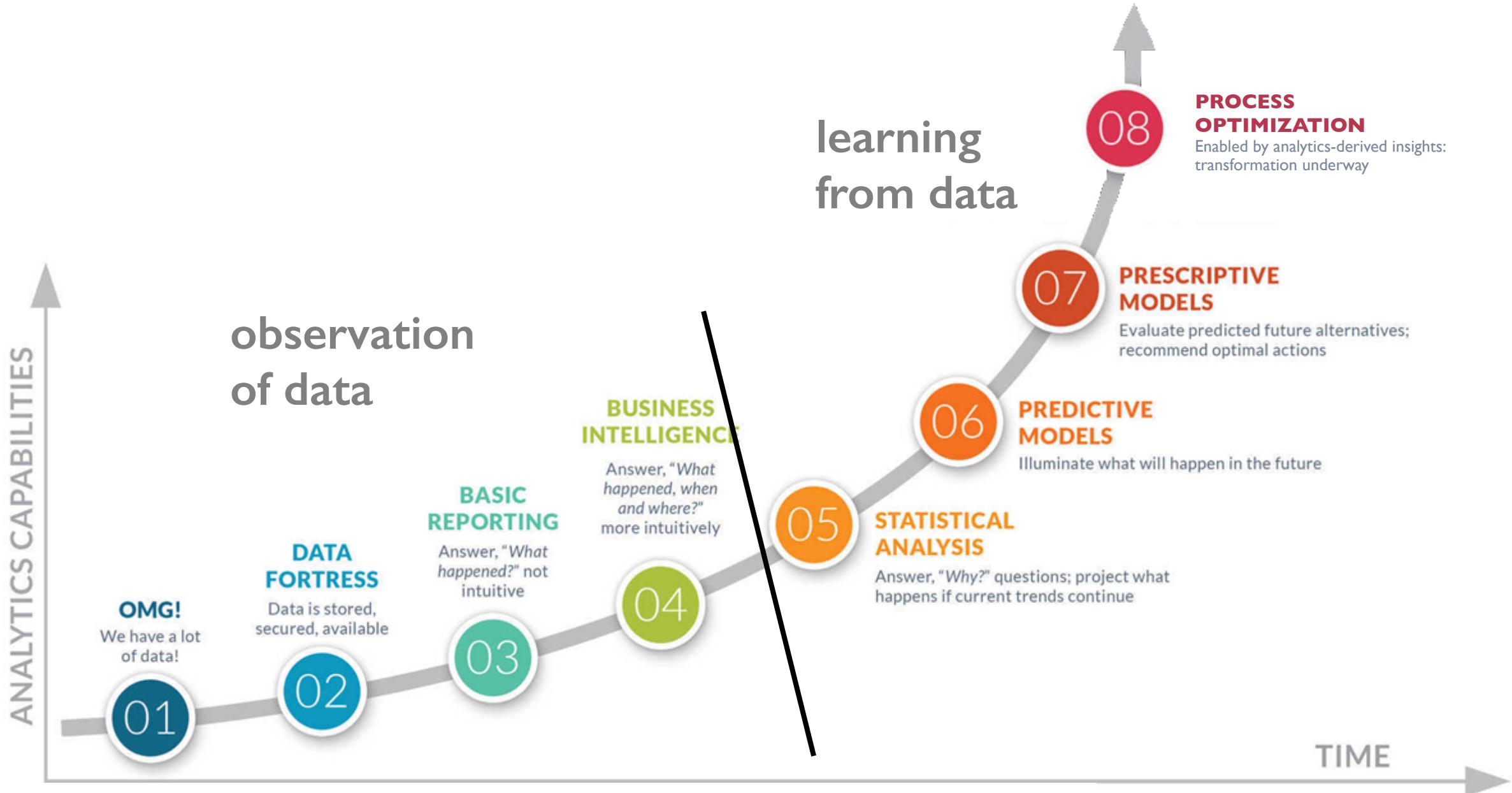
Many of the Games You Know Are Bad

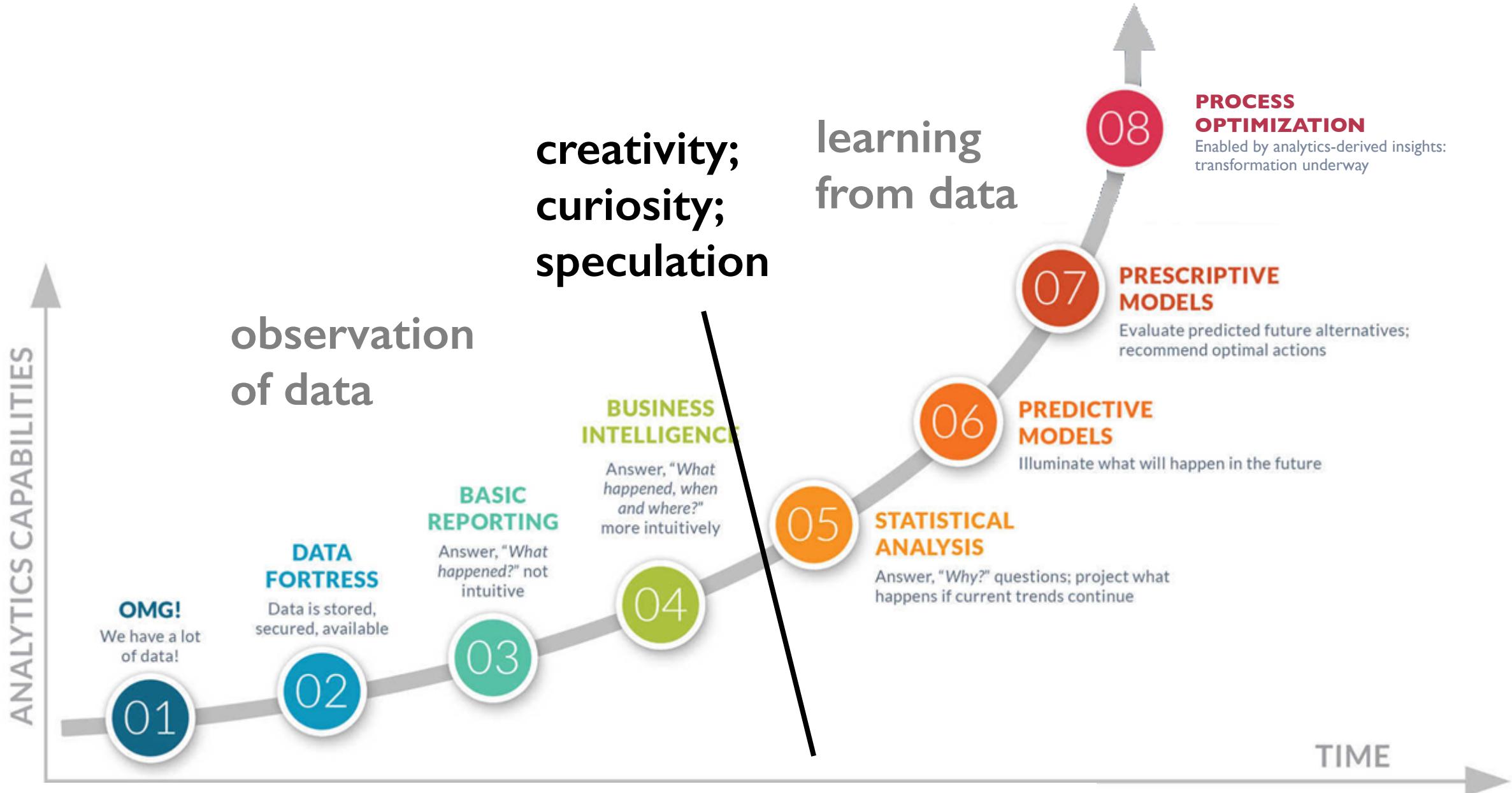
Boardgamegeek average rating vs number of user ratings, filtering to games with at least ten user ratings.



Data from boardgamegeek.com as of 2022-04-17
Analysis at phenrickson.github.io/data-analysis-paralysis/boardgames.html

What is the process producing this data?



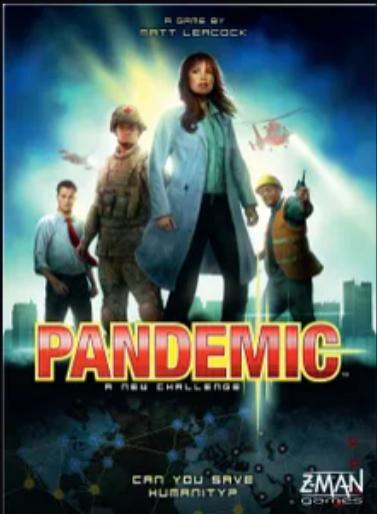


**Can I build models on historical
games to predict which new games
will be highly rated/popular?**

**Can I build models on historical
games to predict which new games
will be highly rated/popular?**

**I might not be able to!
But I can speculate about what
might be useful and then try.**

What explains why games are popular?



REIMPLEMENTED BY: PANDEMIC LEGACY:... + 10 MORE RANK: OVERALL 119 STRATEGY 125 FAMILY 19

7.6 **Pandemic (2008)**
Your team of experts must prevent the world from succumbing to a viral pandemic.

112K Ratings & 18K Comments · GeekBuddy Analysis

2–4 Players Community: 1–4 — Best: 4	45 Min Playing Time	Age: 8+ Community: 10+	Weight: 2.41 / 5 'Complexity' Rating 
--	-------------------------------	----------------------------------	--

Alternate Names: EPIZootic, Pandemic: 10th Anniversary Edition, 19 + [more](#)

Designer: Matt Leacock

Artist: Josh Cappel, Christian Hanisch, Régis Moulun, Chris Quilliams, Tom Thiel

Publisher: Z-Man Games + 34 more

[See Full Credits](#)

Year Released	2008	Mechanisms	Action Points Cooperative Game Hand Management Point to Point Movement Set Collection Trading Variable Player Powers
Designer	Matt Leacock		
Solo Designer	N/A		
Artists	Josh Cappel Christian Hanisch Régis Moulun Chris Quilliams Tom Thiel		

What explains why games are popular?



REIMPLEMENTED BY: PANDEMIC LEGACY:... + 10 MORE RANK: OVERALL 119 STRATEGY 125 FAMILY 19

7.6 **Pandemic (2008)**
Your team of experts must prevent the world from succumbing to a viral pandemic.

112K Ratings & 18K Comments · GeekBuddy Analysis

2-4 Players Community: 1-4 — Best: 4	45 Min Playing Time	Age: 8+ Community: 10+	Weight: 2.41 / 5 'Complexity' Rating 
Alternate Names: EPIZOotic, Pandemic: 10th Anniversary Edition, 19 + more			
Designer: Matt Leacock			
Artist: Josh Cappel, Christian Hanisch, Régis Moulun, Chris Quilliams, Tom Thiel			
Publisher: Z-Man Games + 34 more			

[See Full Credits](#)

Year Released	2008	Mechanisms	Action Points Cooperative Game Hand Management Point to Point Movement Set Collection Trading Variable Player Powers
Designer	Matt Leacock		
Solo Designer	N/A		
Artists	Josh Cappel Christian Hanisch Régis Moulun Chris Quilliams Tom Thiel		

The Modeling Process

Features

Playing Time
Player Count
Publisher
Designer
Artist
Mechanics
Categories

The data generating process.

Outcome(s)

Community Rating

The Modeling Process

Features

Playing Time
Player Count
Publisher
Designer
Artist
Mechanics
Categories

Outcome(s)

Community Rating

The data generating process.

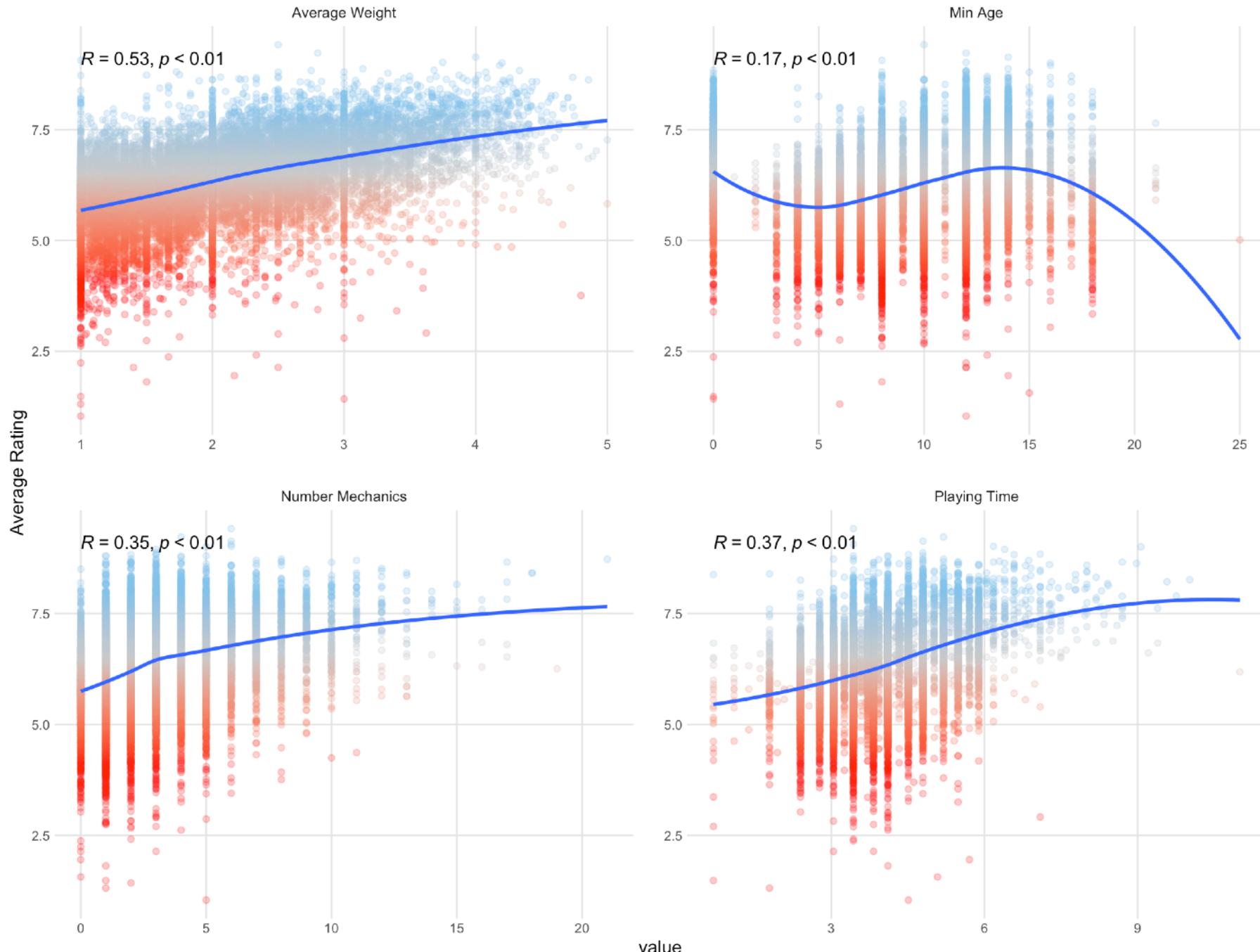
We use a model to learn this.



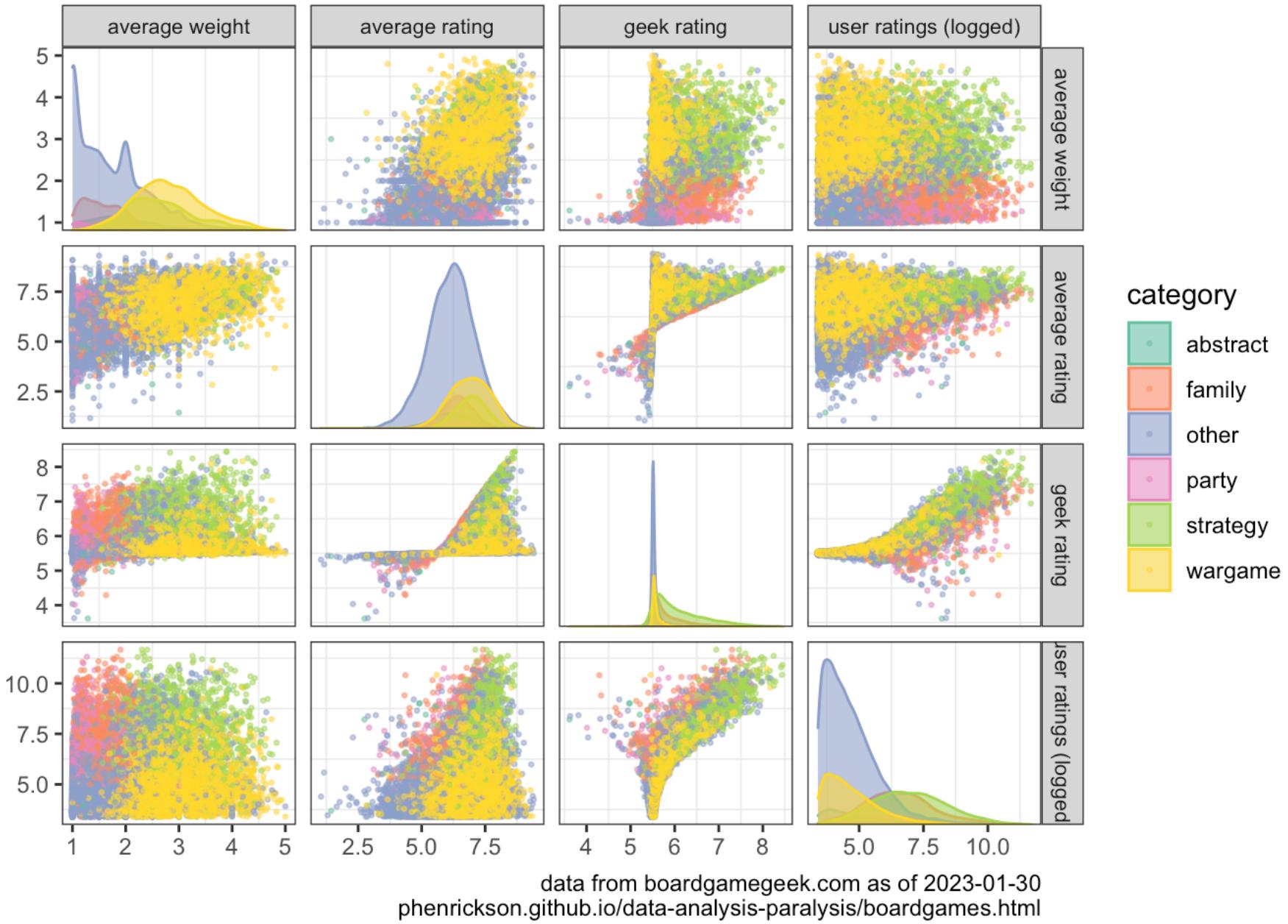
In order to get a model to learn this relationship, we need to **speculate about features** involved in the data generating process.

The data will not do this for us,
and a model can only do so much.

We start looking at the data, and we notice some patterns.

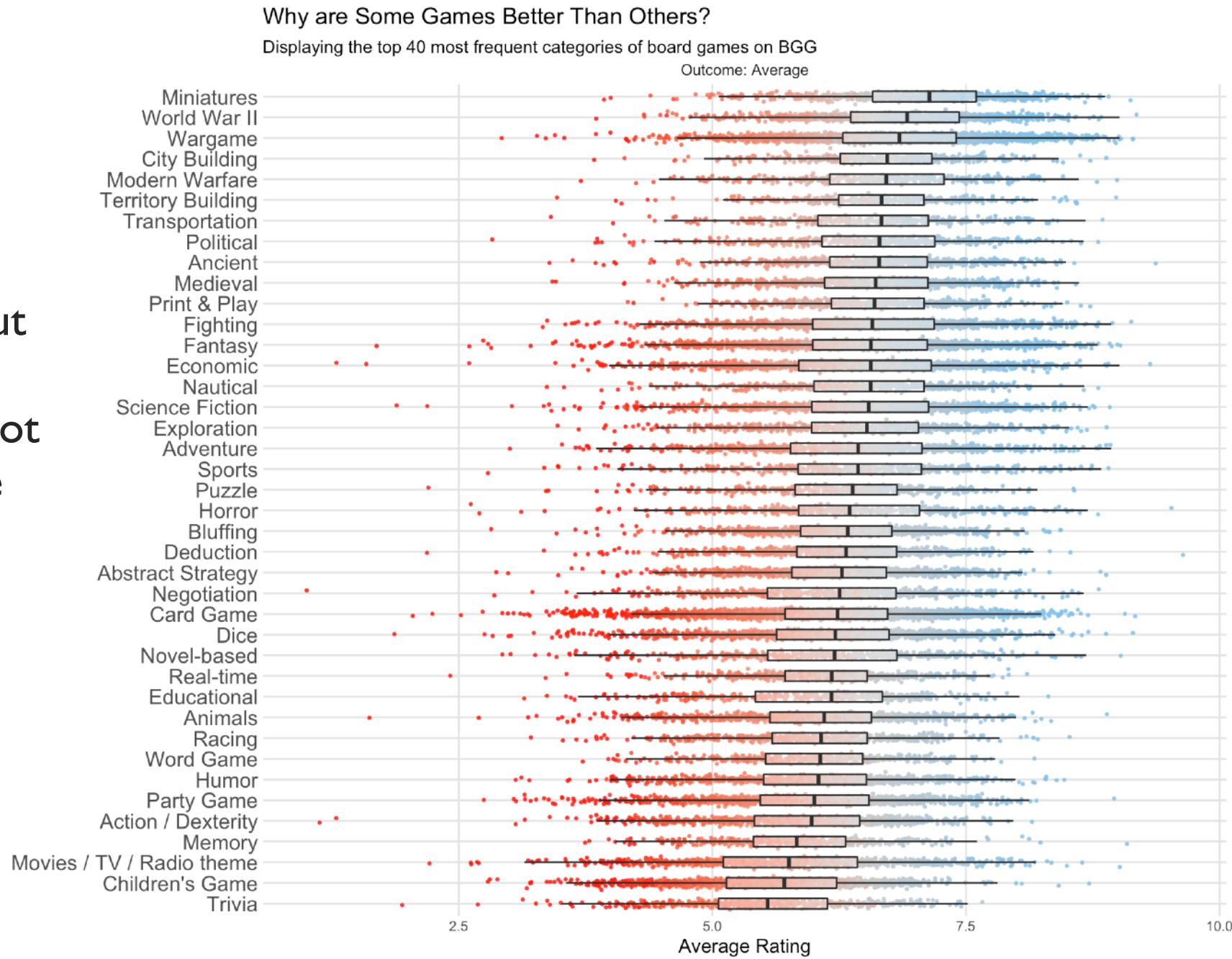


We start looking at the data, and we notice some patterns.



We make decisions about what data to include (and not include) in the model.

Then we train the model.



How does a model learn?

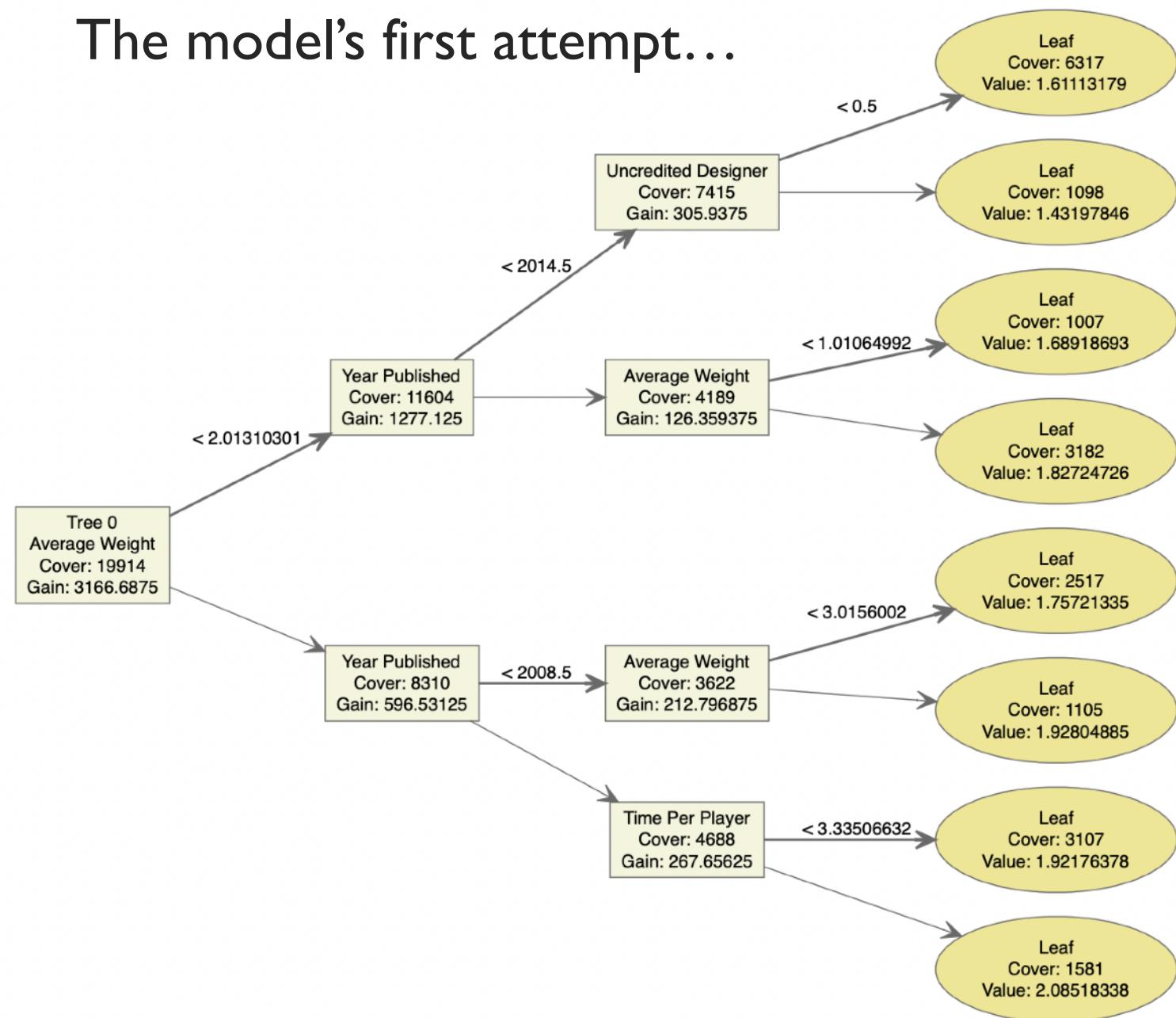
The same way we do. It looks at the data,
finds a pattern, then makes a prediction.
It tries to figure out where it went wrong,
and then it tries again.

How does a model learn?

The same way we do. It looks at the data, finds a pattern, then makes a prediction. It tries to figure out where it went wrong, and then it tries again.

I set the model to look at games published before 2019. Here is the model's first attempt at making predictions.

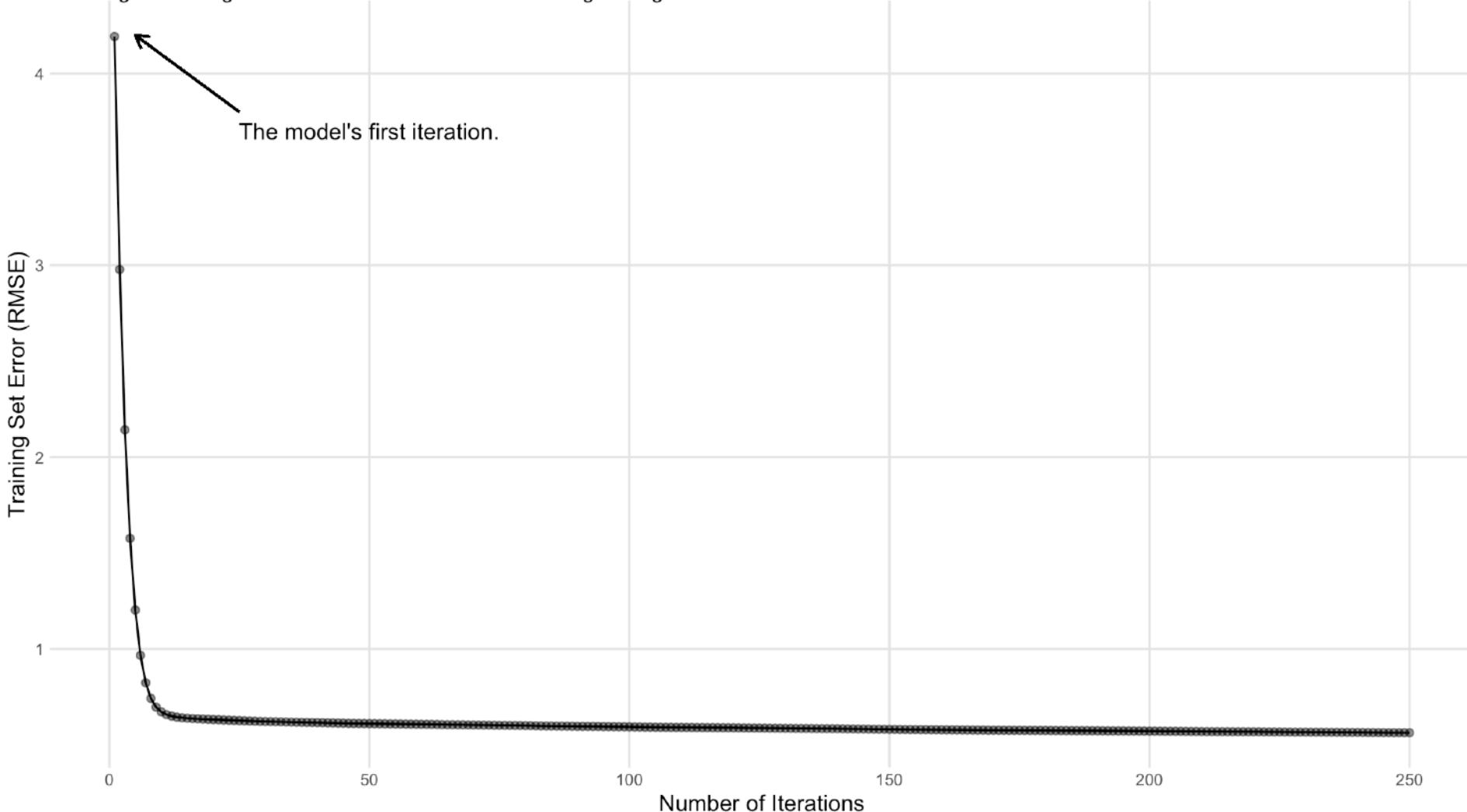
The model's first attempt...



The model's first attempt... **is super inaccurate.**

How Does a Model Learn?

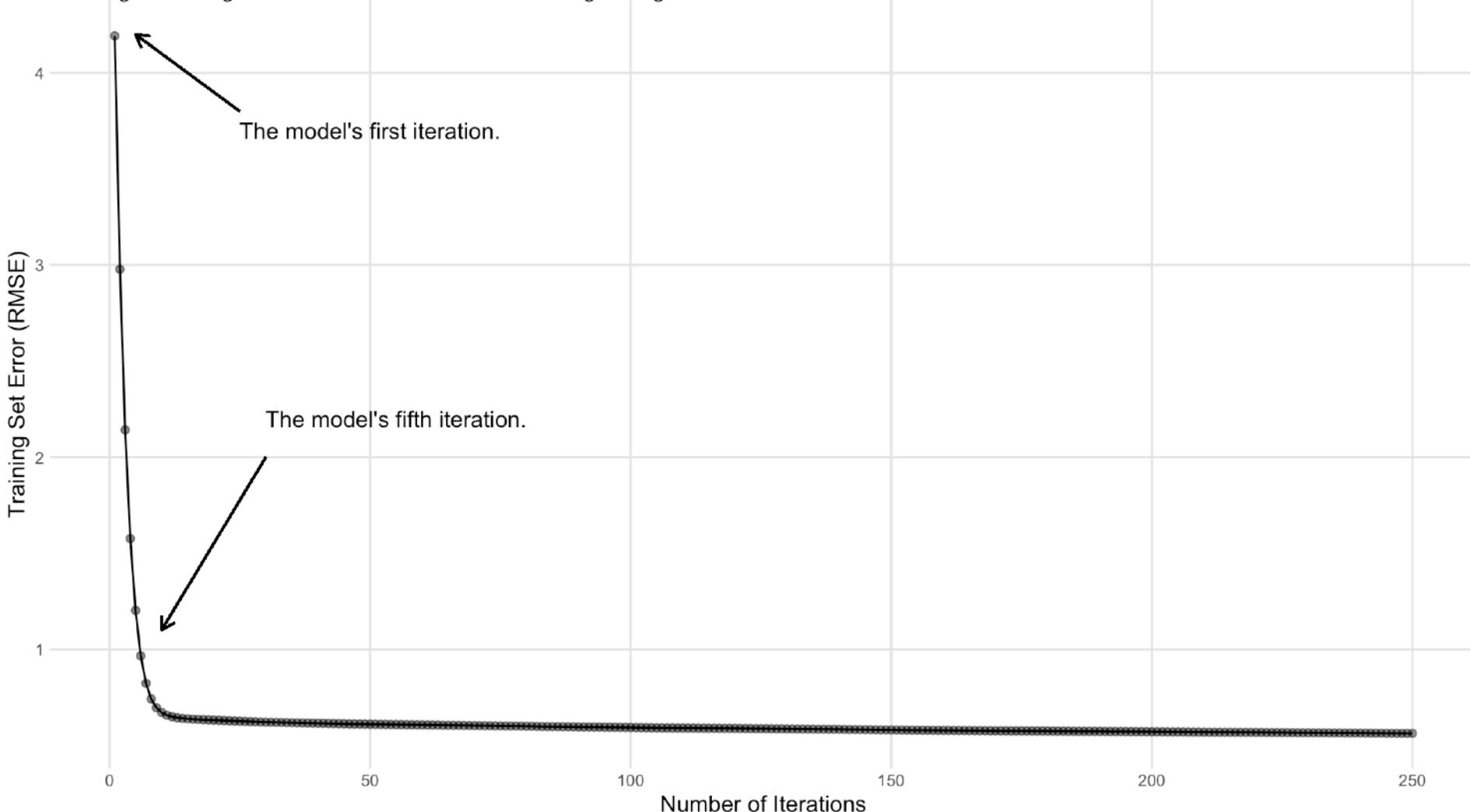
Learning curve for gradient boosted trees trained on average rating.



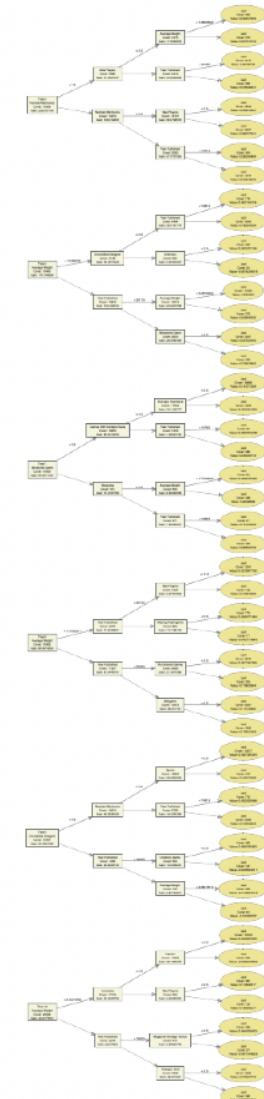
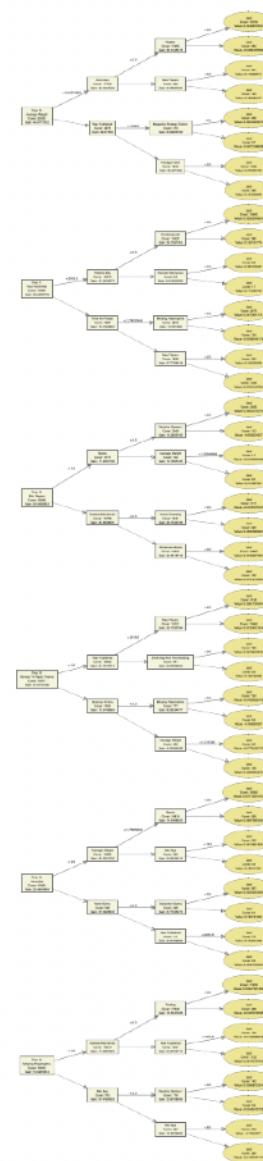
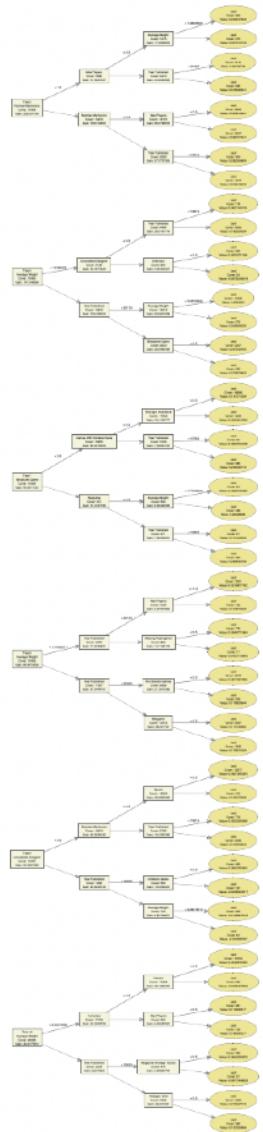
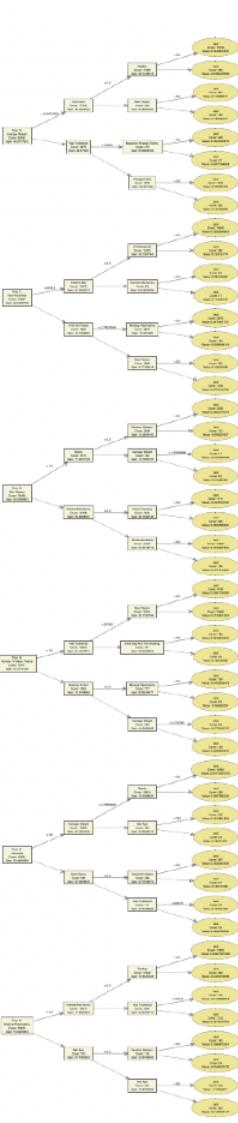
But that's okay! It learns, tries again, and gets better.

How Does a Model Learn?

Learning curve for gradient boosted trees trained on average rating.



The model keeps trying, until I tell it to stop.



At this point comes the test:

**If the model has learned something
about the data generating process,
it should be able to predict new games.**

At this point comes the test:

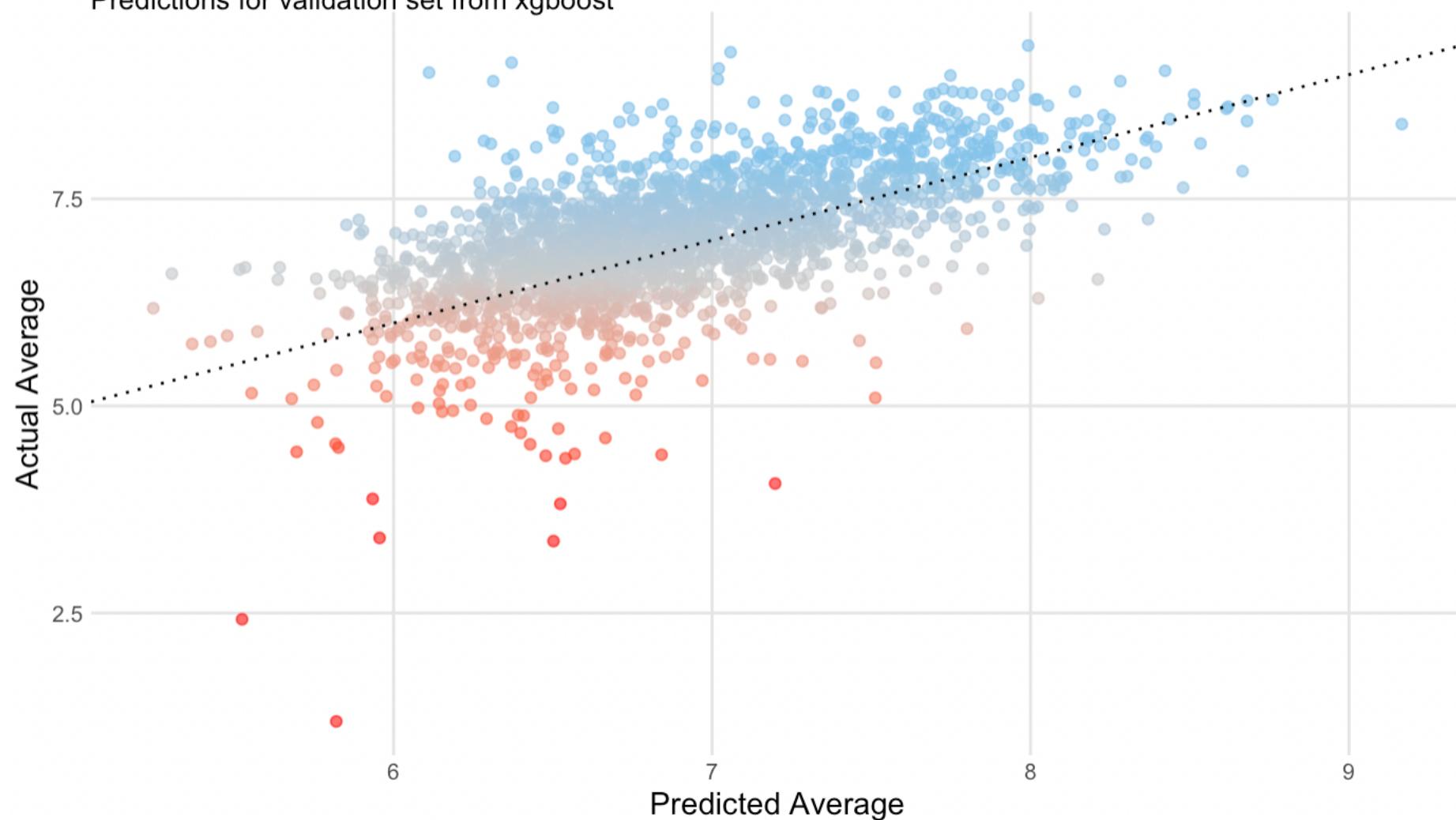
If the model has learned something about the data generating process, it should be able to predict new games.

So, we validate the model by predicting games it hasn't seen before, those released in 2019-2020.

How did it do?

Predicting with a Trained Model

Predictions for validation set from xgboost



The Complex Model
Typical Error: .62

~25% improvement
over simple model

Hooray! The model we trained performed pretty well, better than simpler models.

Are we done? Do we just roll with the model?

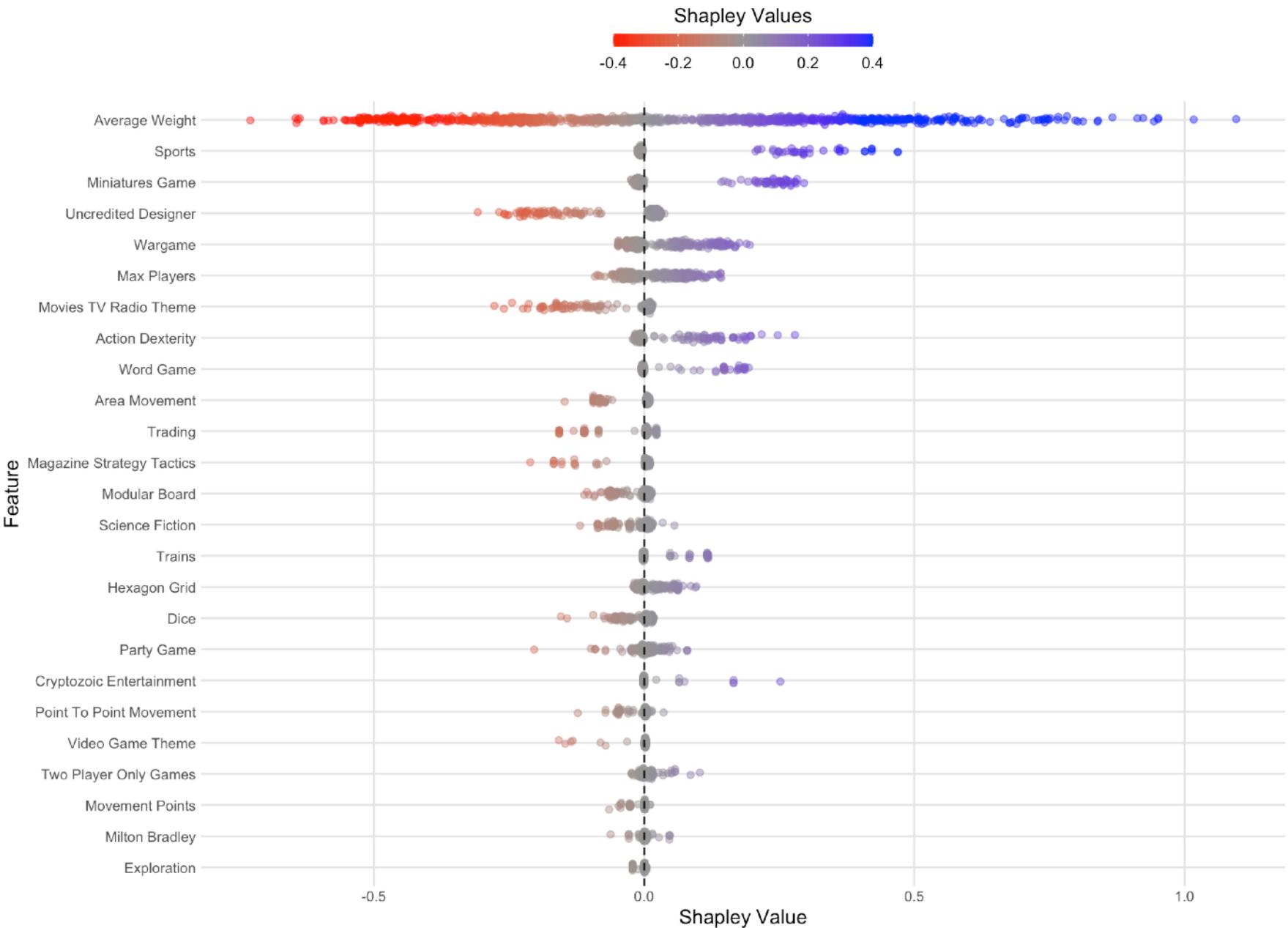
Hooray! The model we trained performed pretty well, better than simpler models.

Are we done? Do we just roll with the model?

No! We need to ask, **what did the model learn? Why is it better?**

What Predict's the Average Rating?

Displaying Shapley values from xgbTree trained on average rating

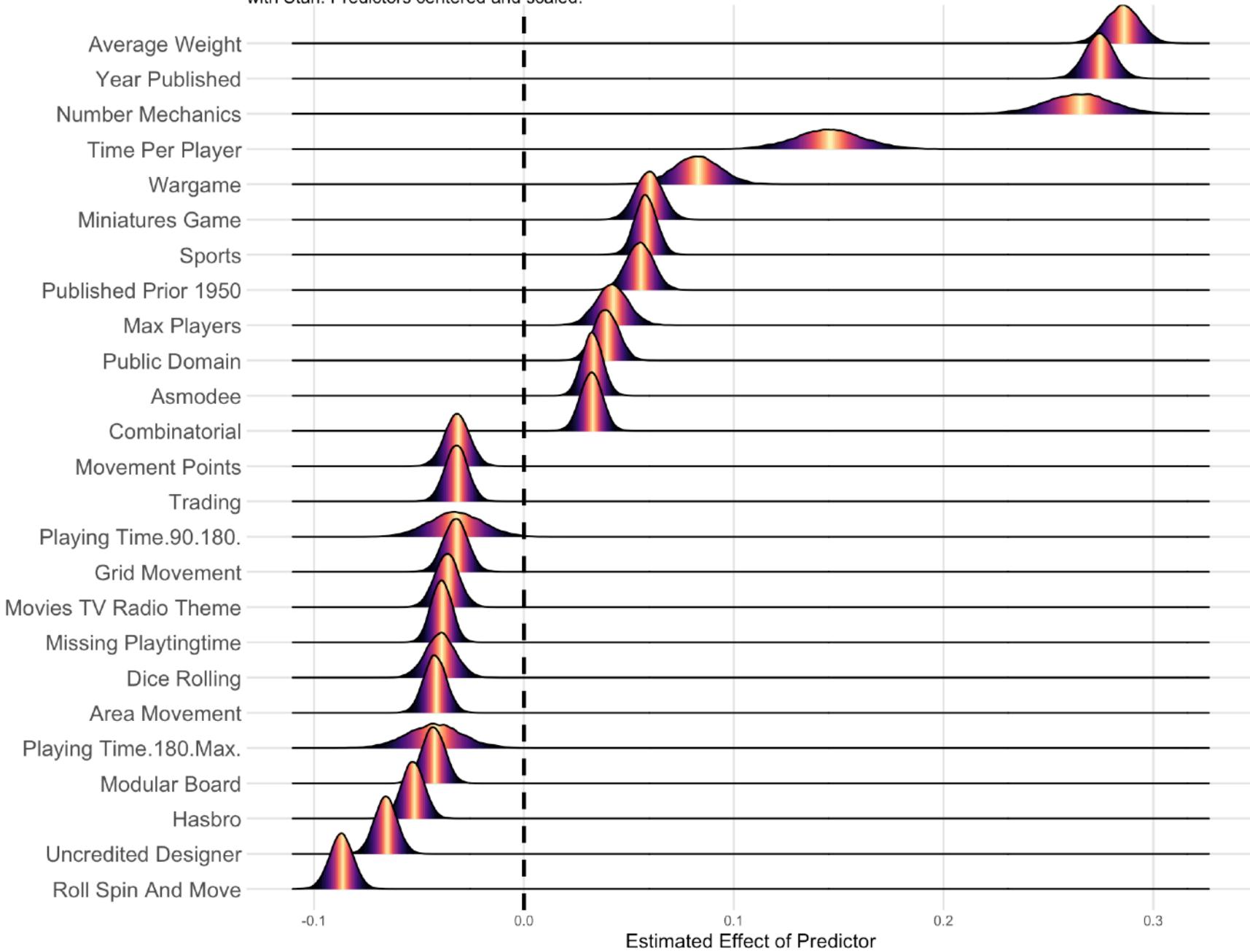


Why are some games rated higher than others?

A model can only tell us what it learned; we have to interpret and understand what this means!

What Predicts the Average Rating?

Displaying estimated coefficients for a linear regression of average rating. Model fit with Stan. Predictors centered and scaled.



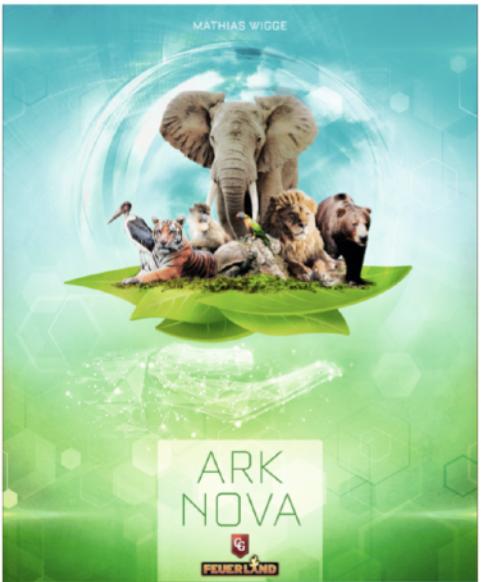
The best games for 2021-2023, according to the model(s).

Rank	Published	ID	Name	Estimated			
				User Ratings	Weight	Average	Geek Rating
1	2022	331106	The Witcher: Old World	26,800	3.60	8.48	8.32
2	2021	343905	Boonlake		4.03	8.18	7.77
3	2021	285967	Ankh: Gods of Egypt		2.81	8.13	7.74
4	2022	310873	Carnegie		3.63	8.36	7.69
5	2022	319807	Shogun no Katana		3.87	8.41	7.66
6	2021	329841	Ticket to Ride: Europe – 15th Anniversary		2.05	8.04	7.62
7	2022	331224	Zombicide: Undead or Alive		2.44	8.06	7.61
8	2021	351735	Newton & Great Discoveries		2.95	8.07	7.61
9	2022	266064	Trudvang Legends		2.54	8.12	7.54
10	2022	302892	Frozen Frontier		4.23	8.33	7.54
11	2021	344277	Corrosion		3.58	7.95	7.54
12	2023	357212	Fire for Light		2.84	8.60	7.54
13	2022	314582	Amsterdam		3.38	7.84	7.51
14	2021	340466	Unfathomable		3.24	7.91	7.49
15	2021	339906	The Hunger		2.03	7.64	7.47
16	2022	349067	The Lord of the Rings: The Card Game – Revised Core Set		2.92	8.00	7.46
17	2022	350198	Terminus		4.21	8.42	7.45
18	2021	342942	Ark Nova		3.65	8.22	7.39
19	2021	322708	Descent: Legends of the Dark		3.27	8.20	7.34
20	2022	359999	Agricola 15		3.54	8.14	7.33
21	2022	295770	Frosthaven		3.64	8.40	7.25
22	2021	332075	Warhammer Quest: Cursed City		3.33	8.29	7.25
23	2022	340520	Ronin Warrior		3.47	8.41	7.20
24	2021	299255	Vienna Connection		2.85	7.85	7.19
25	2021	262201	Sword & Sorcery: Ancient Chronicles		2.97	8.10	7.17

Once we have a model in place, we aren't done. We use it to predict new data and look at what it produces.

See where it does well. See where it makes mistakes.

Some games the model has predicted pretty well.



Ark Nova
ID: 342942
Published: 2021
Player Count: 1-4
Playing Time: 150 min

1.1 Estimated Outcomes on BGG

This table displays the selected game's **current** values on four BGG outcomes (UsersRated, Average, GeekRating, Weight) along with my predictive model(s) **estimated** values for where these games are likely to end up.

Published	ID	Name	Type	UserRatings	Average	GeekRating	Weight
2021	342942	Ark Nova	Current	6,383	8.70	7.94	3.74
			Estimated	3,400	8.22	7.28	3.65

To see more information about the game on boardgamegeek, click on the game's ID or Name to go straight to the game's profile page.

Some
games the
model
hasn't
predicted
so well.



Sleeping Gods
ID: 255984
Published: 2021
Player Count: 1-4
Playing Time: 1200 min

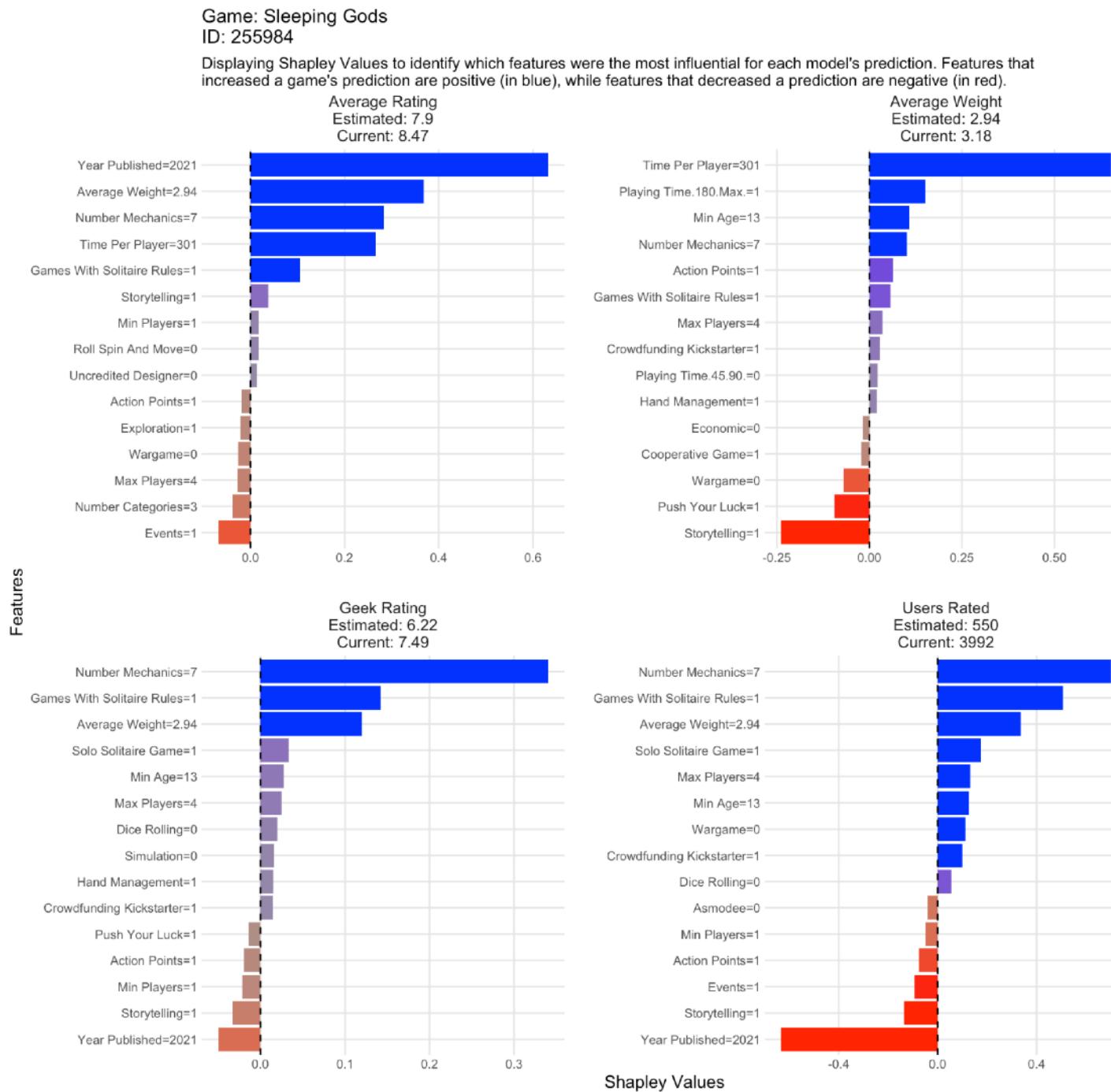
1.1 Estimated Outcomes on BGG

This table displays the selected game's **current** values on four BGG outcomes (UserRatings, Average, GeekRating, Weight) along with my predictive model(s) **estimated** values for where these games are likely to end up.

Published	ID	Name	Type	UserRatings	Average	GeekRating	Weight
2021	255984	Sleeping Gods	Current	3,968	8.47	7.49	3.17
			Estimated	600	7.90	6.10	2.94

To see more information about the game on boardgamegeek, click on the game's ID or Name to go straight to the game's profile page.

Why?
We investigate
the model's
predictions so
that we can
figure out
how to
improve.



Predicted Geek Ratings for Upcoming Games

Estimated ratings for boardgamegeek outcomes from predictive models trained prior to 2022

RANK	PUBLISHED	IMAGE	GAME	DESCRIPTION	COMPLEX	RATINGS	AVERAGE	GEEK
1	2024		Gloomhaven: Second Edition (2024)	Gloomhaven: Second Edition is a revised and elevated version of the award-winning core game of Gloomhaven. This is the culmination of everything Isaac Childres and the growing Cephalofair Games team have learned since the initial release of Gloomhaven, including feedback from the community, playtesters, co-designers, and developers. The world, story, and challenging gameplay are all still the ...	3.8	2600	8.8	7.4
2	2024		Galactic Cruise (2024)	Hello, and welcome to Galactic Cruise. Here, we offer our guests something special: the comfort of a luxury cruise with the innovation of space travel. As the first company to offer extended-stay space vacations, we are excited to have you working for us! As a supervisor of this company, you'll be expected to not only build these shuttles and satisfy our guests, but also to help the company th...	4.2	3750	8.1	7.2
3	2024		Santorini: Pantheon Edition (2024)	NOTE: Pantheon Edition includes all material from the Golden Fleece expansion, as well as every promo card and many new cards. A lot of cards are revamped, reworked and rebalanced. Santorini is an accessible strategy game, simple enough for an elementary school classroom while aiming to provide gameplay depth and content for hardcore gamers to explore. The rules are simple. Each turn consists ...	2.0	3350	7.6	6.8
4	2024		Midhalla (2024)	Midhalla is a cooperative dungeoncrawler that focuses on strategic hero mechanics, tower defense elements, and a character-centered story. The game follows the story of Harkon, a newly trained Viking warrior as he faces the invading demon hordes. At the core of Midhalla's gameplay is a strategic hero combat mechanic that eliminates randomness and focuses on complex strategy. Players invest reso...	3.5	1450	8.5	6.8
5	2024		Zombicide: White Death (2024)	Winter is upon us and its dark and cruel night brings forth dreadful menaces, worse than frostbite... Zombicide Medieval Fantasy is back! With new survivors, new zombies, new setting, and new rules! Scale the great walls to your advantage Use the added height of the fortress walls to escape the Zombies' grasp, dump cauldrons of hot tar onto enemies below, and gain superior line of sight! Ar...	3.0	1800	8.0	6.7
6	2024		1848: The Springtime of Nations (2024)	"Do you not feel, through a kind of instinctive intuition that cannot be analyzed but that is certain, that the ground is shaking once again in Europe? Do you not feel—how shall I call it?—that a revolutionary wind is blowing?" – Alexis de Tocqueville, Souvenir (January, 1848) 33 years after the Congress of Vienna, Continental Europe still lives under the reactionary system designed by the Aus...	3.6	1400	8.3	6.6
7	2024		Dragon Eclipse (2024)	Enter a magical world of mythical beasts and ancient dragons in Dragon Eclipse, the exciting new board game for 1-2 players! Uncover the secrets of the eclipse that changed everything and discover the dangerous and treacherous world that awaits you. Explore a vivid and immersive world filled with majestic and unique creatures you can battle, tame, catch, and train. Each creature has a unique d...	2.8	1100	8.5	6.6
8	2024		A Fading Star: Insurgency and Piracy in Somalia (2024)	A Fading Star: Insurgency and Piracy in Somalia is Volume XV of the COIN Series originally designed by Volko Ruhnke. The year is 2007. We reach the climax of the several decades of civil war that have ravaged Somalia since the fall of the dictator Siad Barre in 1991. An African Union coalition, assisted by the Ethiopian army, enters Mogadishu to support a nascent but corrupt Transitional Federal Go...	3.8	1100	8.4	6.5

I'm on something like my fifth version of the model; I've gathered additional data, tried new techniques.

Predicted Geek Ratings for Upcoming Economic Games

Estimated ratings for boardgamegeek outcomes from predictive models trained prior to 2022

RANK	PUBLISHED	IMAGE	GAME	DESCRIPTION	COMPLEX	RATINGS	AVERAGE	GEEK
1	2024		Galactic Cruise (2024)	Hello, and welcome to Galactic Cruise. Here, we offer our guests something special: the comfort of a luxury cruise with the innovation of space travel. As the first company to offer extended-stay space vacations, we're excited to have you working for us! As a supervisor of this company, you'll be expected to not only build these shuttles and satisfy our guests, but also to help the company th...	4.2	3750	8.1	7.2
2	2024		Midhalla (2024)	Midhalla is a cooperative dungeoncrawler that focuses on strategic hero mechanics, tower defense elements, and a character-centered story. The game follows the story of Harkon, a newly trained Viking warrior as he faces the invading demon hordes. At the core of Midhalla's gameplay is a strategic hero combat mechanic that eliminates randomness and focuses on complex strategy. Players invest reso...	3.5	1450	8.5	6.8
3	2024		Ascendancy (2024)	Ascendancy is the ultimate 4x fantasy board game experience. Lead 1 of many asymmetric houses of royal lineage, each vying for the throne. Reclaim your family name, through a hero's (or villains) journey spanning multiple generations, or a single scenario. Build your empire skyward, seize control outward, manipulate the economy, or dominate the battlefield. There are several paths to victory...	3.8	900	7.9	6.3
4	2024		12 Monks (2024)	This story takes place in the 12th century when Portugal was taking shape... With his armies marching towards the south, Afonso Henriques, the King of Portugal, could not abandon the territory conquered from the Moors and decided to donate a huge plot of land in the center of the country (44,000 hectares) to a religious order of French origin called Cistercians. This community of monks, which ...	3.9	650	7.6	6.0
5	2024		Natera: New Beginning (2024)	In Natera: New Beginning, you play as a sentient and intelligent animal tribe, exploring and controlling areas abandoned in a post-humanity world. With the help of your unique tribe leader and your explorers, you will explore, build authority and take control of 4 distinct areas. Doing so will unlock new, more powerful tiles and allow you to establish settlements to further cement your presenc...	3.3	500	7.7	5.9
6	2024		The Gilded Realms (2024)	The Gilded Realms is a euro-style kingdom builder board game set in a medieval fantasy world, where players will be tasked with ruling over their young but ambitious kingdoms. In this hybrid simultaneous turn-based game of economic management and risk assessment, you and up to three other players will command a unique Faction and compete for the title of ultimate regent. Through careful plann...	3.8	450	7.8	5.9
7	2024		Inferno (2024)	In the Divine Comedy, the poet Dante Alighieri enters Hell in search of his beloved Beatrice. Guided by Virgil, he descends through the nine circles, witnessing how each sinner is eternally punished in a manner as horrifying as the sin they committed in life. Inferno is a soul management game where each player must guide sinners to their respective circles in Hell. The central board consists o...	3.6	350	8.0	5.9
8	2024		House of Fado (2024)	Some Portuguese restaurants are serving typical Portuguese food along with performances of Fado - an iconic world-renowned traditional portuguese music - delivering customers a profoundly expressive style of Portuguese music. People spend the evening in these restaurants just to hear this music. It speaks about "Saudade" (a Portuguese word meaning "longing, nostalgia, yearning, missing somethin...	2.5	500	7.3	5.9

Predicted Geek Ratings for Upcoming War Games

Estimated ratings for boardgamegeek outcomes from predictive models trained prior to 2022

RANK	PUBLISHED	IMAGE	GAME	DESCRIPTION	COMPLEX	RATINGS	AVERAGE	GEEK
1	2024		1848: The Springtime of Nations (2024)	"Do you not feel, through a kind of instinctive intuition that cannot be analyzed but that is certain, that the ground is shaking once again in Europe? Do you not feel—how shall I call it?—that a revolutionary wind is blowing?" - Alexis de Tocqueville, Souvenir (January, 1848) 33 years after the Congress of Vienna, Continental Europe still lives under the reactionary system designed by the Aus..."	3.6	1400	8.3	6.6
2	2024		A Fading Star: Insurgency and Piracy in Somalia (2024)	A Fading Star: Insurgency and Piracy in Somalia is Volume XV of the COIN Series originally designed by Volkol Ruhnke. The year is 2007. We reach the climax of the several decades of civil war that have ravaged Somalia since the fall of dictator Siad Barre in 1991. An African Union coalition, assisted by the Ethiopian army, enters Mogadishu to support a nascent but corrupt Transitional Federal Go...	3.8	1100	8.4	6.5
3	2024		Cysmic (2024)	Cysmic is an epic sci-fi strategy wargame set amidst a crumbling planet. Players assume the role of one of the factions on the planet that is fighting desperately to survive. Throughout the game, players gather resources, capture enemy units, and work to construct their colony ship. Once built, the colony ship must be launched once all of your units are evacuated. The faction to do so first not...	3.6	350	8.2	5.9
4	2024		Night Witches (2024)	Night Witches is a solitaire or two-player cooperative player game in which the Night Witches—female pilots of the Soviet 588th Night Bomber Regiment—work together to fly harassment missions on the Eastern Front during WWII. Their planes are rickety and minimally equipped, but with their engines whirring quietly and with the dark of night to cloak them, the Night Witches are flying towards prom...	2.6	250	8.0	5.8
5	2024		Fate of Venterra (2024)	The War Begins... The mystical maelstrom which protected Venterra through the ages has suddenly vanished. As word spreads nations from around the known world hurry to amass their forces. Legends tell of Venterra's untold wonders and destructive magic, the likes no mortal has ever seen. Fate of Venterra is a war game set in the fictional land of Venterra. A game of pure strategy, with no dic...	3.8	200	8.5	5.8
6	2024		The Republic's Struggle (2024)	The Republic's Struggle is a thematic game with card engine for 2 players, set during the historical events that took place in Spain after the resignation to the throne of King Alfonso XIII, and the proclamation of the II Spanish Republic, in 1931. With the creation of the new regime, the struggle for power between the different social, political and ideological sectors increased, which gave ri...	3.4	200	8.2	5.7
7	2024		After the Last Sky: The First Intifada 1987-1993 (2024)	After the Last Sky simulates the first Palestinian Intifada, from 1987-1993. The game system is a custom implementation of the classic Card Driven Game (CDG) mechanism combined with new mechanisms that simulate civil unrest on a strategic scale. In the game 2 players simulate the First Palestinian Intifada. One side is the Unified National Leadership of the Uprising (UNLU) composed of the Pa...	3.0	200	7.9	5.7
8	2024		Third Crusade (2024)	Unite with your brothers in arms for this holy war. Richard Lionheart and the Angevine have joined with knights from The Holy Roman Empire to try and conquer the city of Jerusalem and retake the holy land from Saladin's Ayyubid dynasty and the Turkomans. Fight for control over the cities and deserts of Palestine. Plan out your actions alongside your ally before turning against them and take th...	3.2	200	7.9	5.7

Building one model tends to generate new ideas for other models, or new data to collect.

At some point I wondered, could I build a model to predict games that I like?

What's the data generating process?



The Modeling Process

Features

Playing Time
Player Count
Publisher
Designer
Artist
Mechanics
Categories

Outcome(s)

Me
deciding
to buy
games

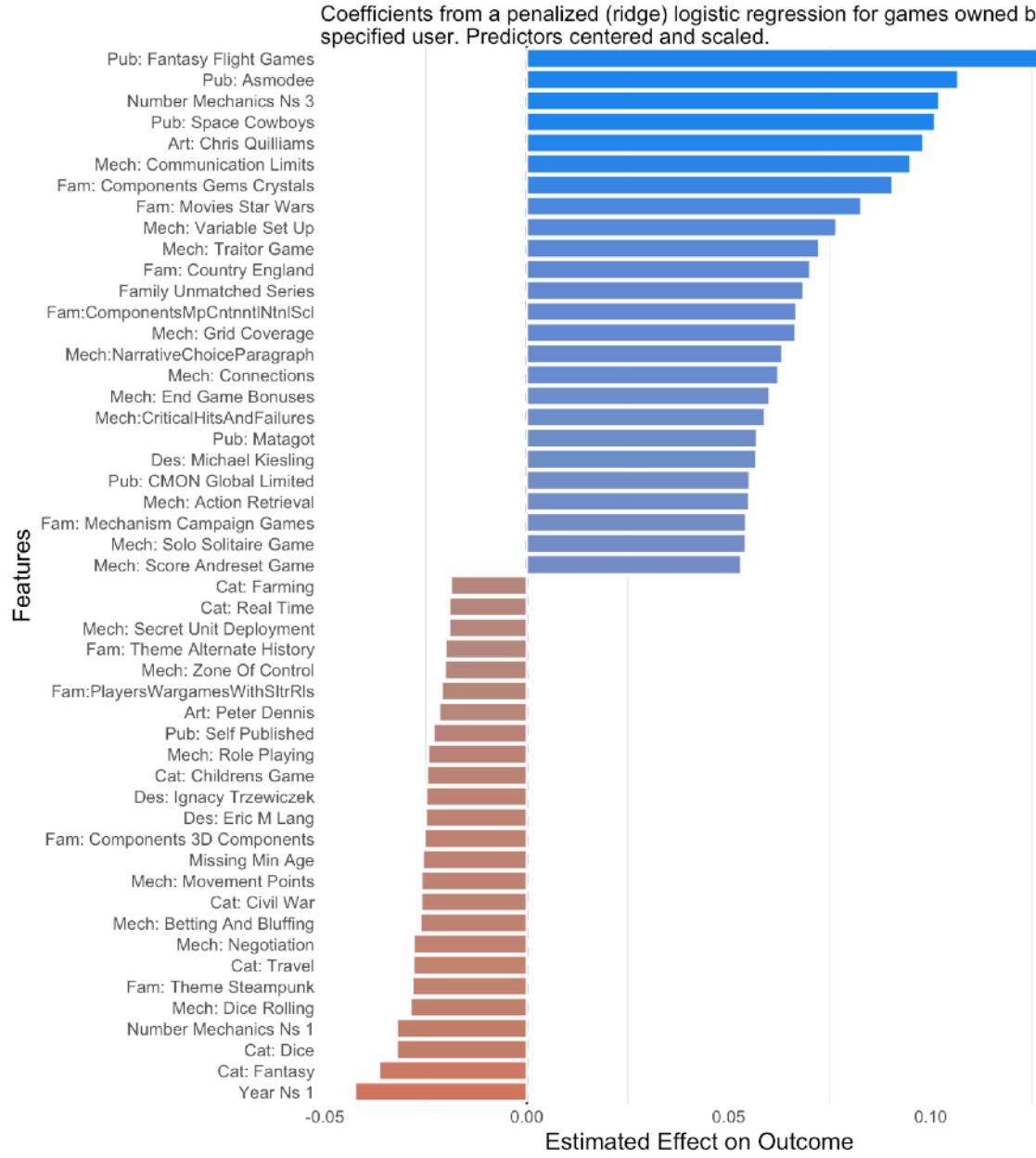
Is the Game in My
Collection



Can a model
learn this?

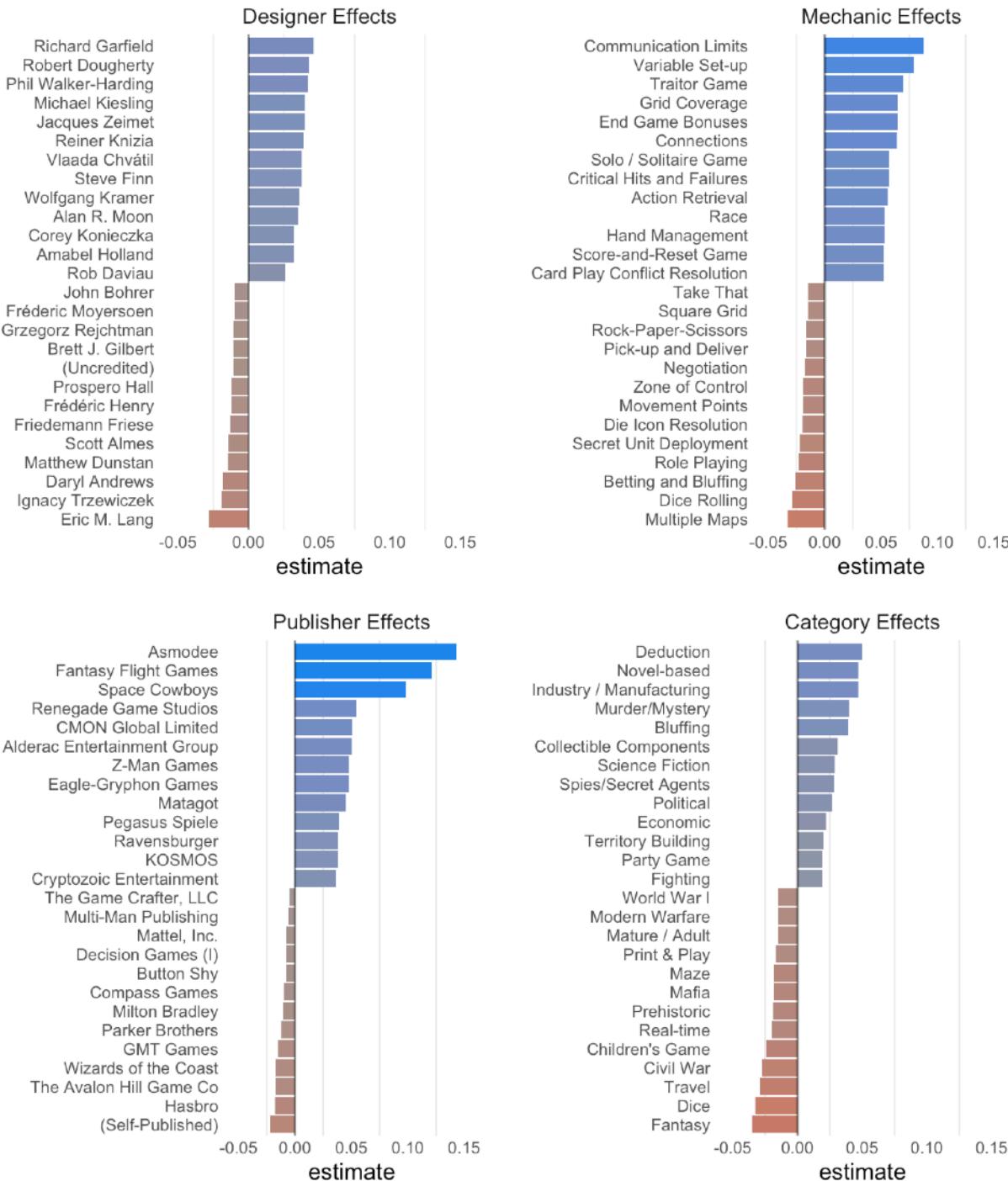
Artificial Phil:

The model of myself and what it learned about what makes me likely to own a game



Artificial Phil:

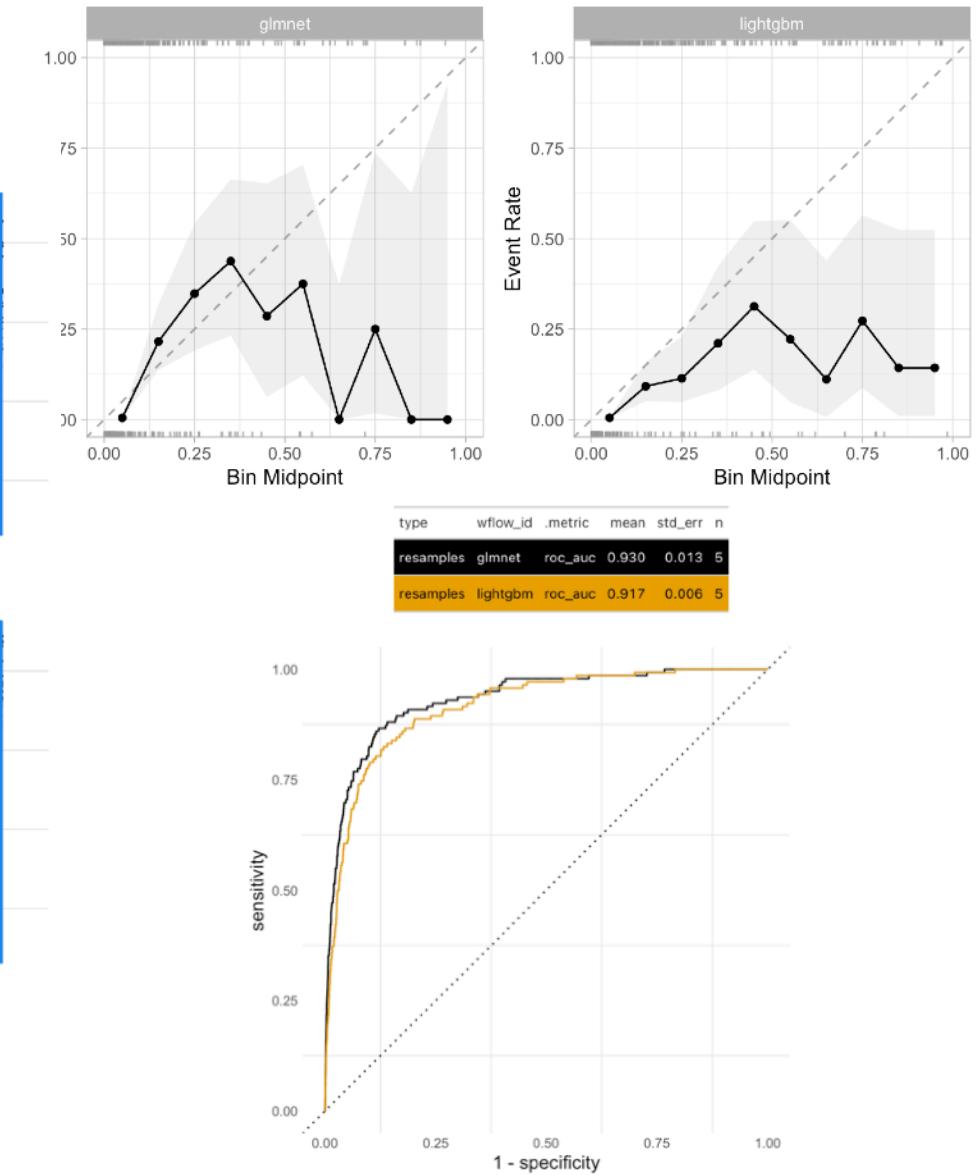
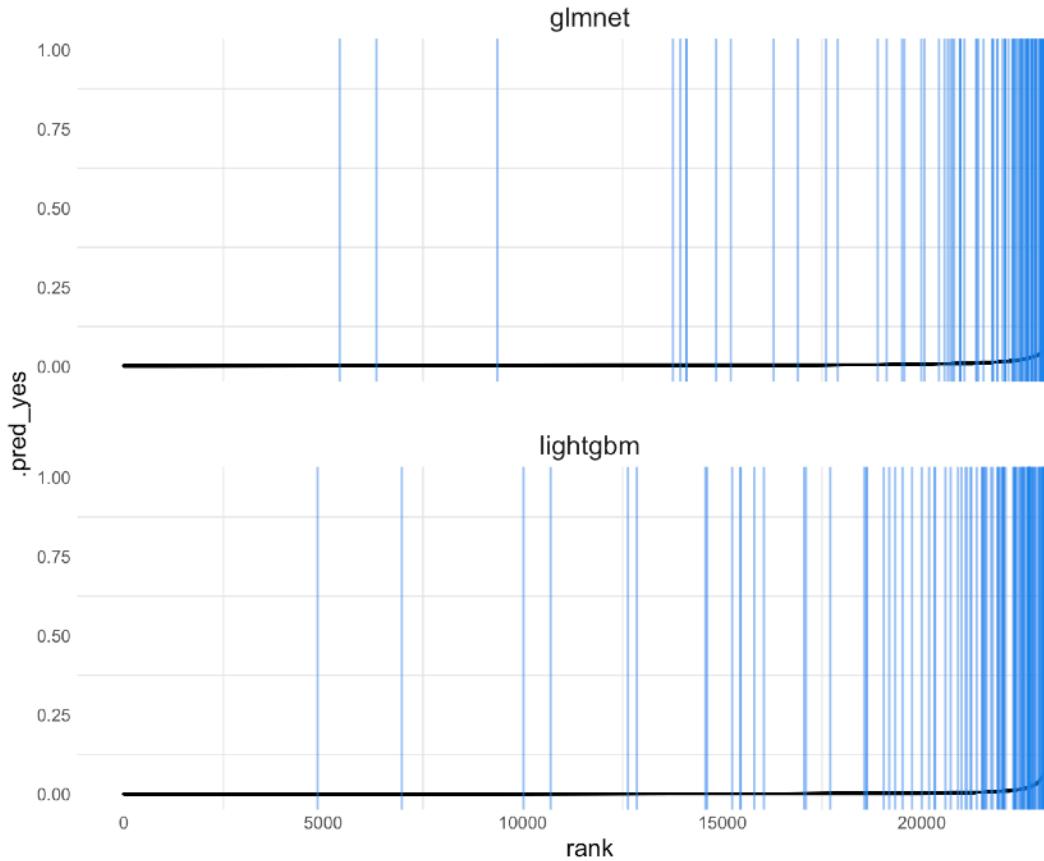
The model of myself and what it learned about what makes me likely to own a game



How well Artificial Phil did at predicting myself

How well did the models do?

Displaying cross validated probabilities for all games in the training set from least likely to most likely. Vertical blue lines indicate game was actually in the user's collection.



Top (Older) Games

Rankings based on predictive model trained on user's collection using games released through 2021

RANK	IMAGE	GAME	DESCRIPTION	PR(OWN)	OWN
1		Twilight Imperium: Fourth Edition (2017)	Twilight Imperium (Fourth Edition) is a game of galactic conquest in which three to six players take on the role of one of seventeen factions vying for galactic domination through military might, political maneuvering, and economic bargaining. Every faction offers a completely different play experience, from the wormhole-hopping Ghosts of Creuss to the Emirates of Hacan, masters of trade and ec...	0.985	yes
2		Unmatched: Little Red Riding Hood vs. Beowulf (2020)	In battle, there are no equals. ONCE UPON A TIME, Little Red Riding Hood took her basket of nasty tricks and faced off against the legendary Beowulf in this exciting Unmatched set. "What big eyes you have, Wulfe!" "That's called 'rage', kid!" Little Red features a clever card-combo mechanism. Matching icons on the cards she plays to the one in her "basket" (discard pile), triggers potent e...	0.969	no
3		Cosmic Encounter Duel (2020)	The Cosmic Citizenship Council has announced it will allow two new alien species to join its ranks, but they forgot to make two copies of the filing form — which means that only one species can join! Now, the two candidates must battle for control of the planets to determine who deserves the right to become a Certified Civilization. Cosmic Encounter Duel is a competitive standalone two-player ...	0.969	no
4		Arkham Horror: The Card Game (Revised Edition) (2021)	The boundaries between worlds have drawn perilously thin. Dark forces work in the shadows and call upon unspeakable horrors, strange happenings are discovered all throughout the city of Arkham, Massachusetts, and behind it all an Ancient One manipulates everything from beyond the veil. It is time to revisit that which started it all... With a revamped system of organization and a number of quali...	0.966	no
5		Newton (2018)	The middle of the 17th century was a period of great changes; with the advent of the scientific method came what we now call the Scientific Revolution. Many great scientists, with their theories and ideas, changed and shaped our perception of the universe: Galileo Galilei, Copernicus, Kepler, Bacon and, above all, Sir Isaac Newton. In Newton, the players take the role of a young scientist who ...	0.965	no
6		Concordia Venus (2018)	Concordia Venus is a standalone reimplementation of Concordia with some added features. Concordia Venus is a peaceful strategy game of economic development in Roman times for 2-6 players aged 13 and up. Instead of looking to the luck of dice or cards, players must rely on their strategic abilities. Be sure to watch your rivals to determine which goals they are pursuing and where you can outpac...	0.954	no

Older games
Artificial Phil
thought I was
most likely to
own



Twilight Imperium: Fourth Edition
Released in 2017
Has a Space Lion on the Cover



You and your friends bring a warring universe to life on your kitchen table

Takes roughly 8-12 hours to play

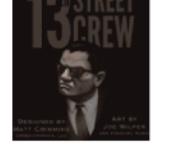
Top 15 (Newer) Games for mrbananagrabber

Rankings based on predictive model trained on user's collection using games released through 2021

RANK	IMAGE	GAME	DESCRIPTION	PR(OWN)	OWN
1		Gloomhaven: Second Edition (2024)	Gloomhaven: Second Edition is a revised and elevated version of the award-winning core game of Gloomhaven. This is the culmination of everything Isaac Childres and the growing Cephalofair Games team have learned since the initial release of Gloomhaven, including feedback from the community, playtesters, co-designers, and developers. The world, story, and challenging gameplay are all still the ...	0.726	no
2		Unmatched: Jurassic Park - Dr. Sattler vs. T. Rex (2022)	In battle, there are no equals. "Dinosaurs eat man... Woman inherits the earth." The greatest predator the world has ever known is closing in on the tenacious Dr. Sattler. Who has the slightest idea what to expect? In Unmatched: Jurassic Park - Dr. Sattler vs. T. Rex, the massive T Rex unleashes fearsome attacks and seems unstoppable while Dr. Sattler makes full use of her surroundings and the a...	0.684	yes
3		Ticket to Ride Legacy: Legends of the West (2023)	In Ticket to Ride Legacy: Legends of the West, players embark on twelve journeys across North America as 19th century pioneers. The campaign begins on the East Coast, with players working their way to the West from one adventure to the next, meeting challenges along the way. As in Ticket to Ride, completing your tickets will remain your primary goal, but you will need to develop other skills if...	0.678	no
4		Terminus (2023)	You and your competitors' transit companies have been hired by the city to build new subway lines and commercial developments to improve the city's bottom line. Manage assets such as time, money, & resources to build your subway line. Gain prestige by completing objectives and fulfilling the city's transit demands. Focus on individual projects, open Agendas or a little of both in an effort t...	0.543	no
5		Circadians: Chaos Order (2022)	The initial quakes were only minor tremors, but as the land began to unravel, so did our sense of security. We watched the cliffs of Hytazch fall into the sea. Mighty trees of old, swallowed up by the caverns below. As the waters rose, a great roar was heard across the plains. This was no cry of disbelief or heartache, but of jubilance. Songs began to fill the air as our once peaceful hosts, no...	0.540	no
6		Captain's Log (2022)	What is the Captain's Log the board game? It is a 1-4 player sandbox board game with an estimated playing time of 1-4 hours and recommended for people aged 14+ where you will be in charge of a ship from the colonial period and you will compete against other players to become the most famous captain of all. The game starts with the selection of our ship. You will have a choice between a swift ...	0.486	no
7		The Lord of the Rings: The Card Game - Revised Core Set (2022)	Sometimes, in order to truly appreciate a tale, one must first go back to its beginning. Grand adventures and strong fellowships are important and wonderful, but the first step of any journey is just as important as the last. With that in mind, it's time to return to the beginning of one of the most epic adventures of all... With increased contents and some quality-of-life improvements, this new...	0.438	yes
		Nucleum	When Elsa von Frühlingfeld presented her invention to King Frederik		

New games
Artificial
Phil thinks
I'm likely to
buy

New games Artificial Phil thinks I'm likely to buy

9		Unmatched: Teen Spirit (2023)	Unmatched is a highly asymmetrical miniature fighting game for two or four players. Each hero is represented by a unique deck designed to evoke their style and legend. Tactical movement and no-luck combat resolution create a unique play experience that rewards expertise, but just when you've mastered one set, new heroes arrive to provide all new match-ups. Unmatched: Teen Spirit features four ...	0.393	no
10		Unmatched: For King and Country (2023)	Unmatched is a highly asymmetrical miniature fighting game for two or four players. Each hero is represented by a unique deck designed to evoke their style and legend. Tactical movement and no-luck combat resolution create a unique play experience that rewards expertise, but just when you've mastered one set, new heroes arrive to provide all new match-ups. Unmatched: For King and Country featu...	0.393	no
11		Frosthaven (2022)	Frosthaven is the story of a small outpost far to the north of the capital city of White Oak, an outpost barely surviving the harsh weather as well as invasions from forces both known and unknown. There, a group of mercenaries at the end of their rope will help bring back this settlement from the edge of destruction. Not only will they have to deal with the harsh elements, but there are other, ...	0.343	yes
12		Unmatched: Houdini vs. The Genie (2022)	Unmatched is a highly asymmetrical miniature fighting game for two or four players. Each hero is represented by a unique deck designed to evoke their style and legend. Tactical movement and no-luck combat resolution create a unique play experience that rewards expertise, but just when you've mastered one set, new heroes arrive to provide all new match-ups. Unmatched: Houdini vs. The Genie adds ...	0.326	no
13		The 13th Street Crew (2023)	The 13th Street Crew is a semi-cooperative social deduction game of criminal strategy. The players are low-ranking members of a large criminal organization headed by the Old Don that for all intents and purposes runs this fair city. The players represent fellow crew members occupying the lowest rung in the organization. Most of the players are ambitious and eager to prove they deserve to adv...	0.322	no
14		Quatro City (2022)	Quatro City is a one-of-a-kind wooden puzzle with unique pieces and a catching detective quest to be solved. Each piece with a detailed and bright illustration is soaked with the mysteries. This puzzle brings leisure activities to a whole new level by adding interaction and activity to routine puzzle assembling. The first challenge facing you is to assemble the puzzle of the numerous streets o...	0.308	no
15		Cartagena: Escape Diaries (2023)	Cartagena: Escape Diaries is based on the classic game Cartagena and features multiple ways to play. In the original game, now dubbed the "First Escape", each player has a group of six animal pirates, and you want to be the first to have all six escape through the tortuous underground passage that connects the fortress to the port, where a sloop is waiting for them. To move a pirate, you need...	0.305	no

We **use models** in order to
~~predict upcoming games so that we can~~
~~buy them before other people~~
learn about the world around us.

How do we learn from data?
How do we build models?

We speculate.
We test.
We learn from our mistakes.
Science proceeds with trial and error.

Or, as someone very wise once said:

Or, as someone very wise once said:

Pass on what you have learned. Strength.
Mastery. But weakness, folly, failure also.

Or, as someone very wise once said:

Pass on what you have learned. Strength.
Mastery. But weakness, folly, failure also.

Yes, failure most of all.
The greatest teacher, failure is.

Or, as someone very wise once said:

Pass on what you have learned. Strength.
Mastery. But weakness, folly, failure also.

Yoda,
Jedi Master

Yes, failure most of all.
The greatest teacher, failure is.



Wrapping up.

Now What?

Model Building, Science, and Analytics

What should we do?

I'll leave you with one thought.

I'll leave you with one thought.

The answer isn't from computer
science, statistics, mathematics, or
the social sciences.

I'll leave you with one thought.

**The answer isn't from computer
science, statistics, mathematics, or
the social sciences.**

**The answer, naturally, comes from
television.**

The Good Place

**Mike Schur,
Writer**



What does it mean to be a good person?

We can run the full gamut on this, and explore every possible theory about how to be a good person, and it starts to get exhausting.

It is asking too much of people to become monks and shed all of their earthly possessions.

**Mike Schur,
Writer**



You know what's important?
If you're trying.

If you're just trying to be a good person... if you're asking yourself, am I doing okay, could I be improving somehow?

If you're just asking the questions, that's kind of the key.

Mike Schur,
Writer



Try to be a little bit better today
than you were yesterday.

**Mike Schur,
Writer**



It's more important to be a good person than a good scientist.

But, we can easily amend this for the journey into data science.

**Mike Schur,
Writer**



Phil Henrickson, Stealer of Quotes



What does it mean to be a good scientist?

We can run the full gamut on this, and **explore every possible theory** about how to be a good scientist, and it starts to get exhausting.

It is asking too much of people to...
become grad students and shed all of their earthly possessions.

**Phil Henrickson,
Stealer of Quotes**



You know what's important?
If you're trying.

If you're just trying to learn about the world around you... if you're asking yourself, am I doing okay, could I be improving somehow?

**If you're just asking the questions,
that's kind of the key.**

Phil Henrickson, Stealer of Quotes



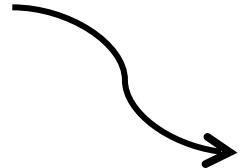
Phil Henrickson, Stealer of Quotes

Try to be a little bit better today
than you were yesterday.



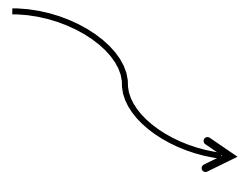
**Try to be a little bit better today
than you were yesterday.**

Try to be a little bit better today
than you were yesterday.



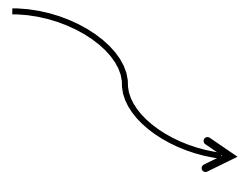
This is the secret to building models.

Try to be a little bit better today
than you were yesterday.



This is the secret to **building models**.
This is the secret to **good science**.

Try to be a little bit better today
than you were yesterday.



This is the secret to building models.

This is the secret to good science.

This is the answer to the now what of
analytics.

Thanks for listening.

Questions?

Appendix

Prerequisites for a data science project

Two Things You Need for a Data Science Project

- 1) There is some unknown pattern that would be useful to learn.
- 2) There is data to learn that pattern.

A model of the model building process:

Data is used here.

1) **We observe.** We notice a pattern or result that has occurred in the world.

2) **We speculate.** We develop explanation for the process that could have produced our observation.

3) **We develop implications.** We ask, if our speculation is correct, what else should we expect to observe?

4) **We test.** We look to see whether the other implications of our model are supported in the data.

The science,
**how we learn
from data,** is
here.