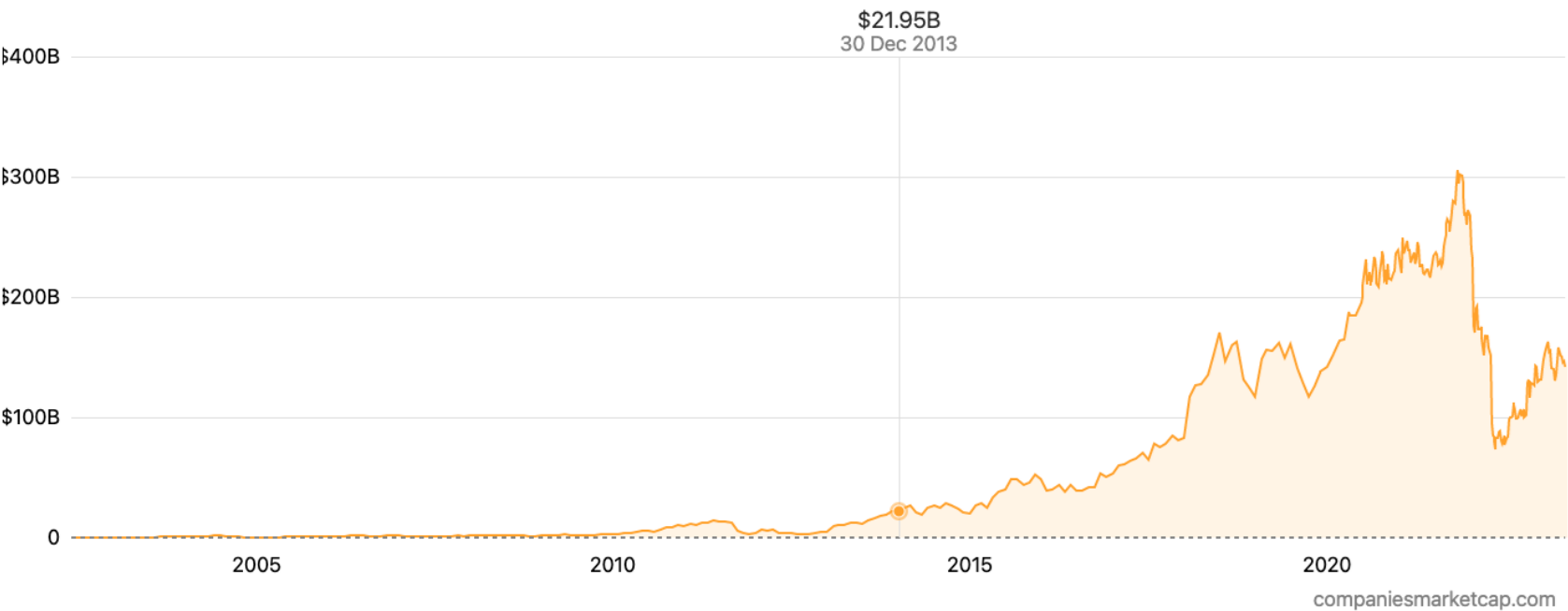# LLO 8200: Introduction- What is Data Science, Anyway?

# Data Science Success: Netflix

**Market cap history of Netflix from 2002 to 2023**

$21.95B
30 Dec 2013

- $400B
- $300B
- $200B
- $100B
- 0

2005     2010     2015     2020

companiesmarketcap.com

# How Does Netflix Use Data?



**Analytics can take many forms at Netflix**

How different analytics personas on the Data Science and Engineering team spend their time

OCCASIONAL CONTRIBUTIONS ▢▢▢▢ PRIMARY FOCUS

| The Analyst | The Engineer | The Visualizer |
| --- | --- | --- |
| Software engineering | Software engineering | Software engineering |
| Infrastructure and tools | Infrastructure and tools | Infrastructure and tools |
| Data engineering | Data engineering | Data engineering |
| Data modeling | Data modeling | Data modeling |
| Analytic tools | Analytic tools | Analytic tools |
| Deep dive analysis | Deep dive analysis | Deep dive analysis |
| Metric development | Metric development | Metric development |
| Statistical modeling | Statistical modeling | Statistical modeling |

**NETFLIX**
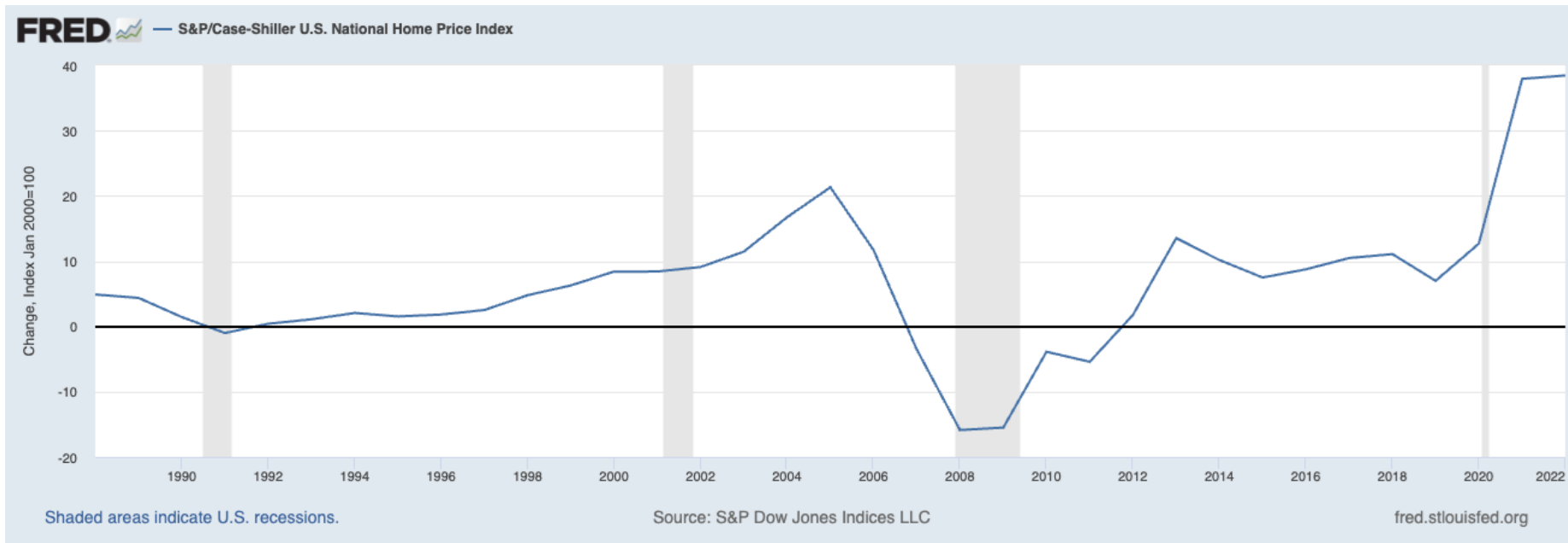
# What Happened to Companies That Didn't Use Modeling?

# Data Science Failure: Zillow



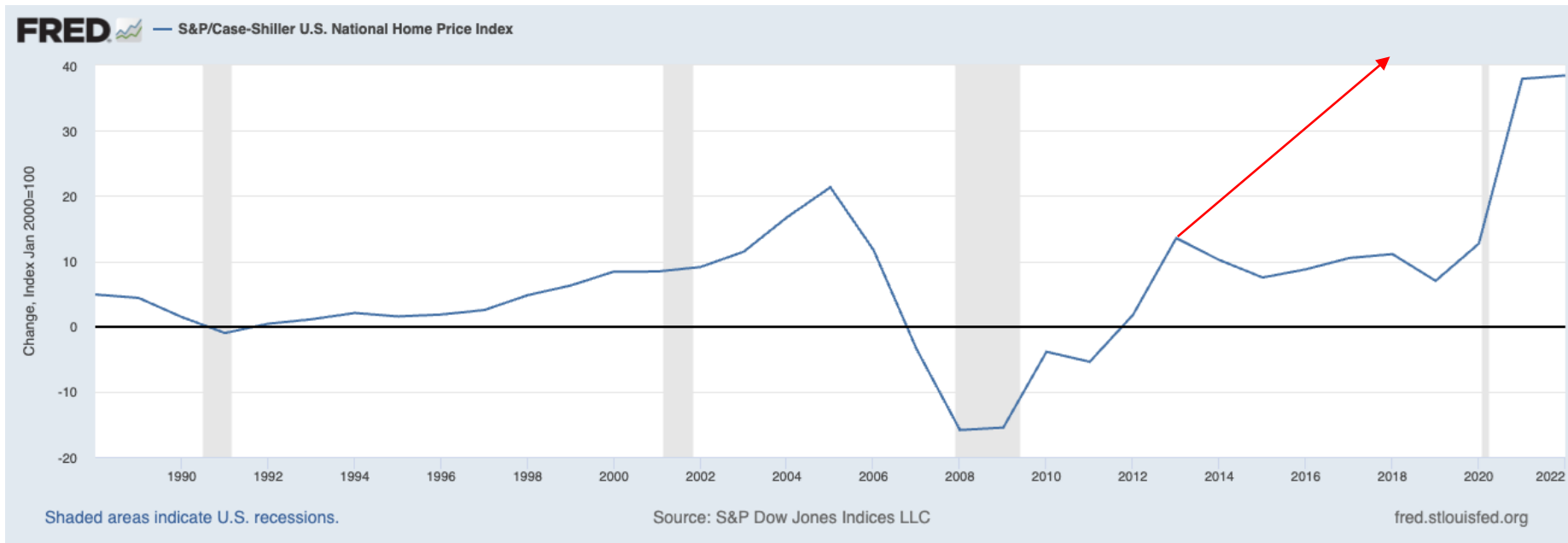**Zillow, facing big losses, quits flipping houses and will lay off a quarter of its staff.**

The real estate website had been relying on its algorithm that estimates home values to buy and resell homes. That part of its business lost about $420 million in three months.
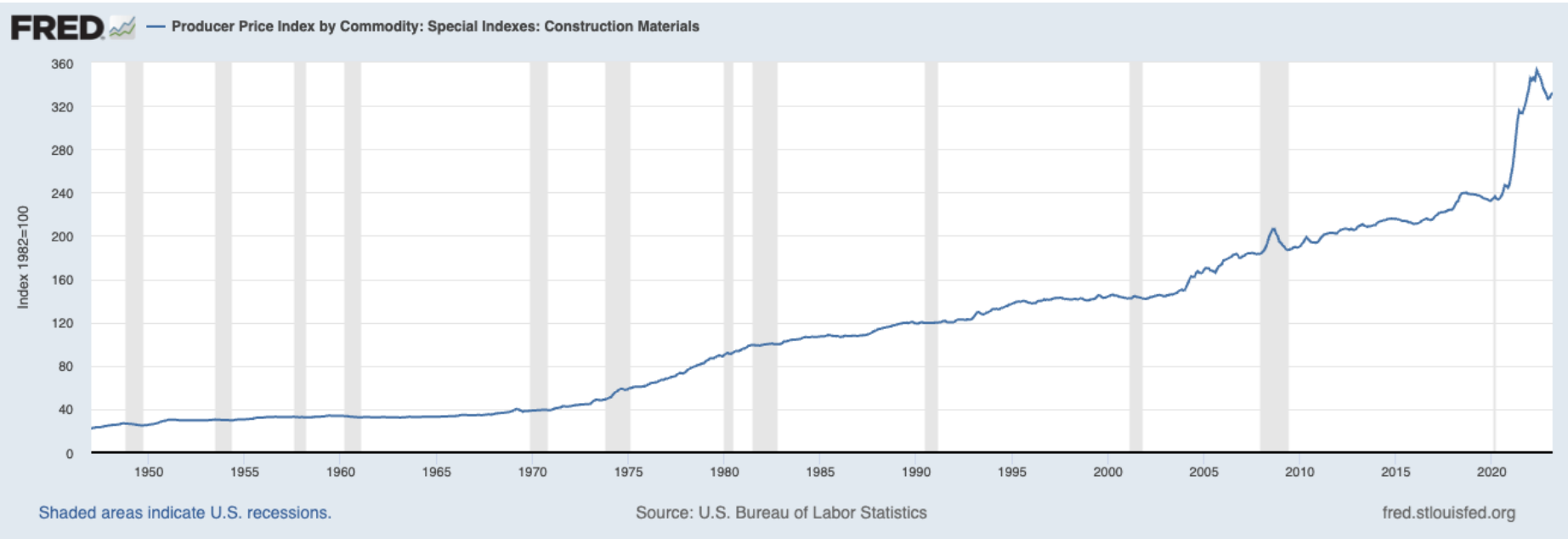
# Data Science Failure: Zillow

# Data Science Failure: Zillow
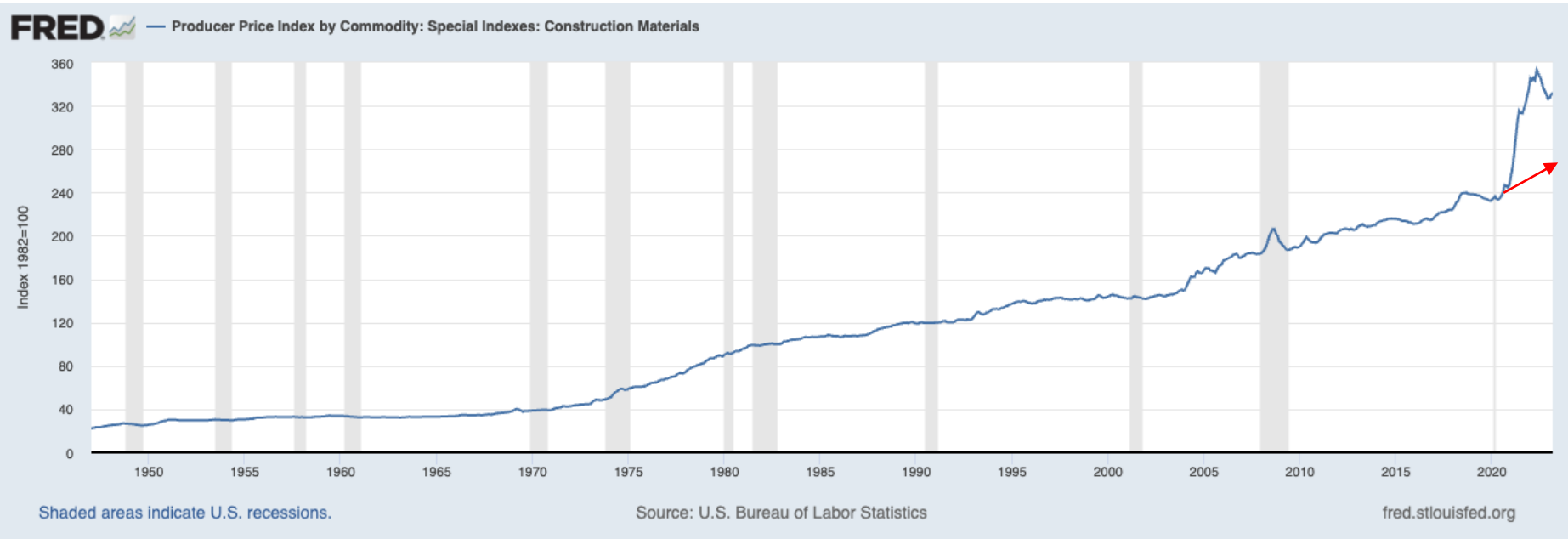
# Data Science Failure: Zillow



FRED — Producer Price Index by Commodity: Special Indexes: Construction Materials

Shaded areas indicate U.S. recessions.

Source: U.S. Bureau of Labor Statistics

fred.stlouisfed.org

# Data Science Failure: Zillow



FRED — Producer Price Index by Commodity: Special Indexes: Construction Materials

Shaded areas indicate U.S. recessions.

Source: U.S. Bureau of Labor Statistics

fred.stlouisfed.org

# Data Science Failure: Zillow

"Data models, … are not good at things that have not happened before," says Aleksandar Tomic, associate dean for strategy, innovation, and technology at Boston College. "A computer will do whatever you ask it to do, but [the outcome] depends on what you ask."

# Some Terms

- Dependent Variable or Outcome: What we're interested in, the measurable phenomena that can change as a result of actions

- Independent Variable or Predictor: Anything that might help to predict the outcome

- Model: A way of establishing the relationship between the predictors and the outcome

# So, What Does Data Science Do?

Descriptives: establish the nature of the dependent variable--

What's the measure of central tendency?

Where is the outcome highest or lowest?

When was the outcome highest or lowest?

How does the outcome vary as a function of other "stuff" that's out there

Modeling: establish relationships between the dependent variable and predictors

How much does the dependent variable increase as a function of the independent variables?

How well does that relationship hold up? Are there times when the link between the outcome and the predictors doesn't seem as strong?

Can we derive actionable insights from the results of the model?

# More Reading

https://towardsdatascience.com/zillow-collapse-proves-ai-is-not-perfect-let-me-explain-3bdeb03b7713

https://fortune.com/education/articles/what-zillows-failed-algorithm-means-for-the-future-of-data-science/

https://netflixtechblog.com/analytics-at-netflix-who-we-are-and-what-we-do-7d9c08fe6965

https://towardsdatascience.com/deep-dive-into-netflixs-recommender-system-341806ae3b48