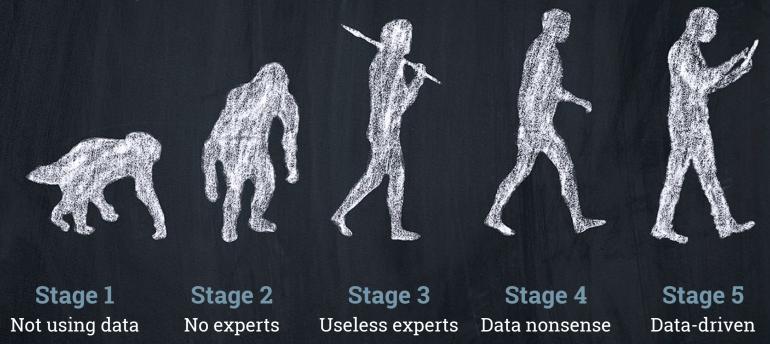
DATA-DRIVEN CULTURE

Cassie Kozyrkov, Chief Decision Scientist, Google

Evolving a culture





Moving past Stage 1: Not using data





Some of it is easy, some of it is hard.

Get started on the easy bits.

You might not need the hard bits.

Everyone is qualified to look at data. There's plenty of free software that makes it easy. Enjoy!







Share data where possible.

The ideal data-driven culture defaults to data openness within its workforce wherever data is not sensitive. Some datasets must be closed for user privacy and/or legal reasons, but you also have data that is just about your organization (e.g. hardware tests, amount of paper printed, org charts, etc.) - why not share that internally?



Inspiration is cheap, rigor is expensive.

Sticking to the data is easy, going beyond it takes expertise.



A map of data science

Get inspired

Make important decision

Many small decisions

Q



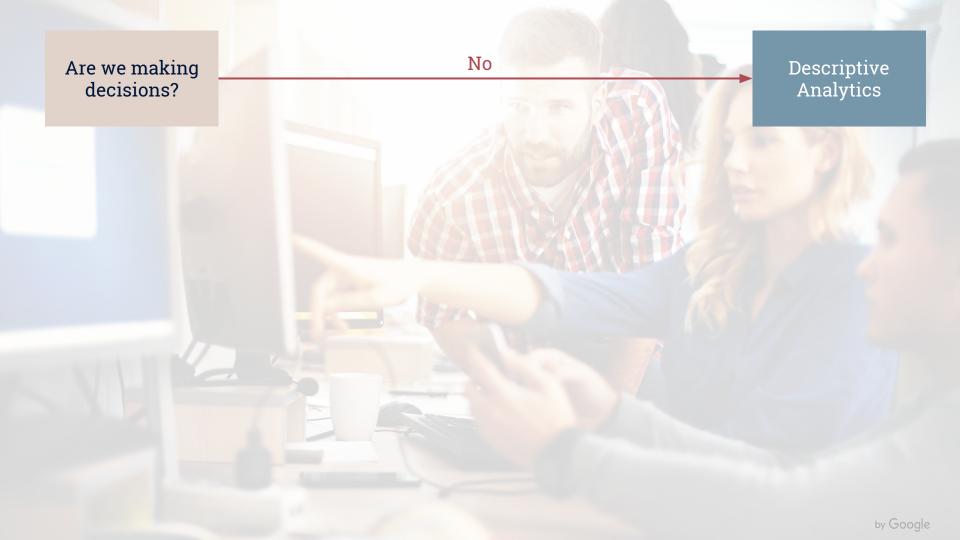


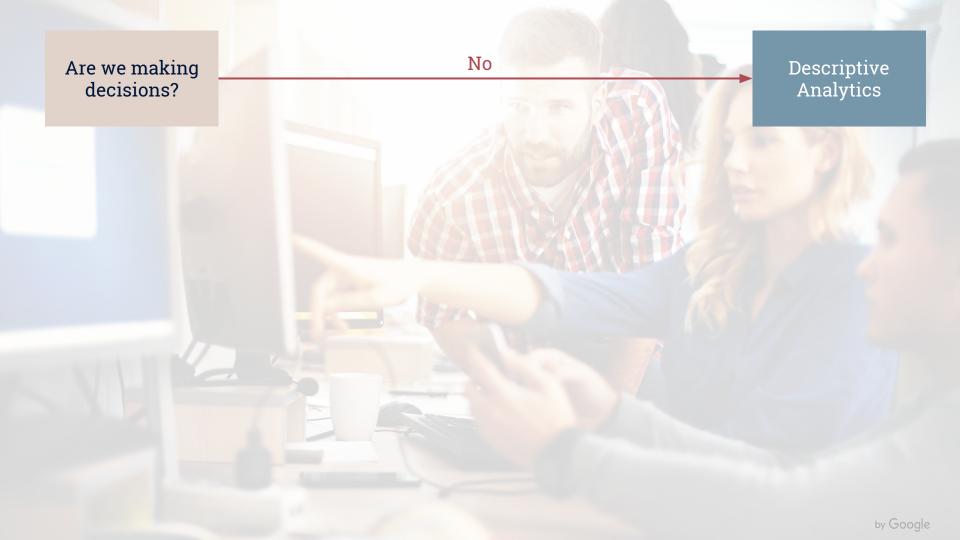
Descriptive Analytics

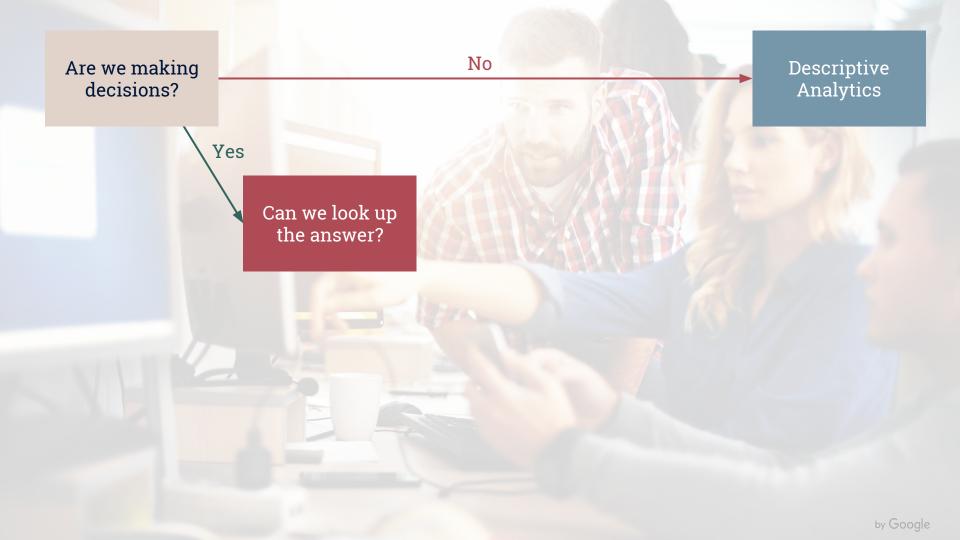
Statistical Inference

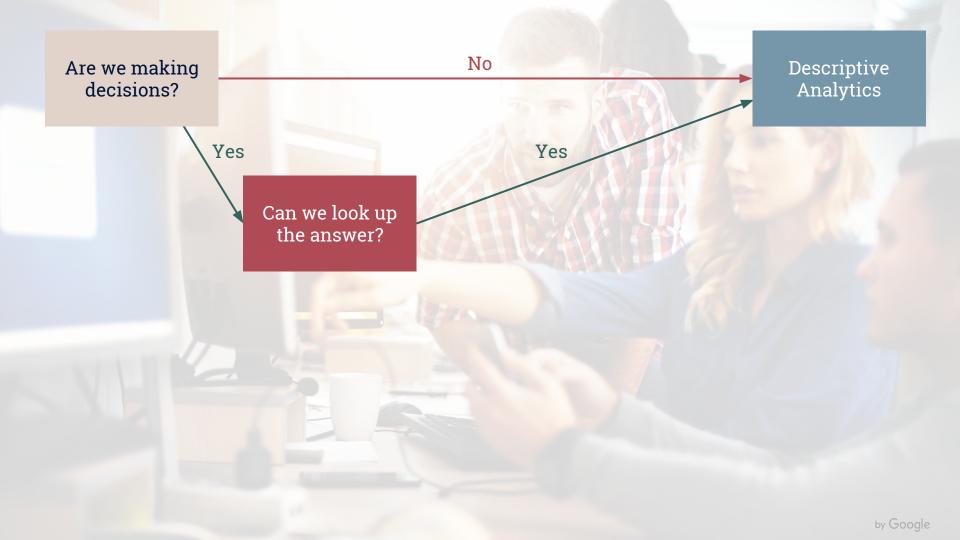
Machine Learning

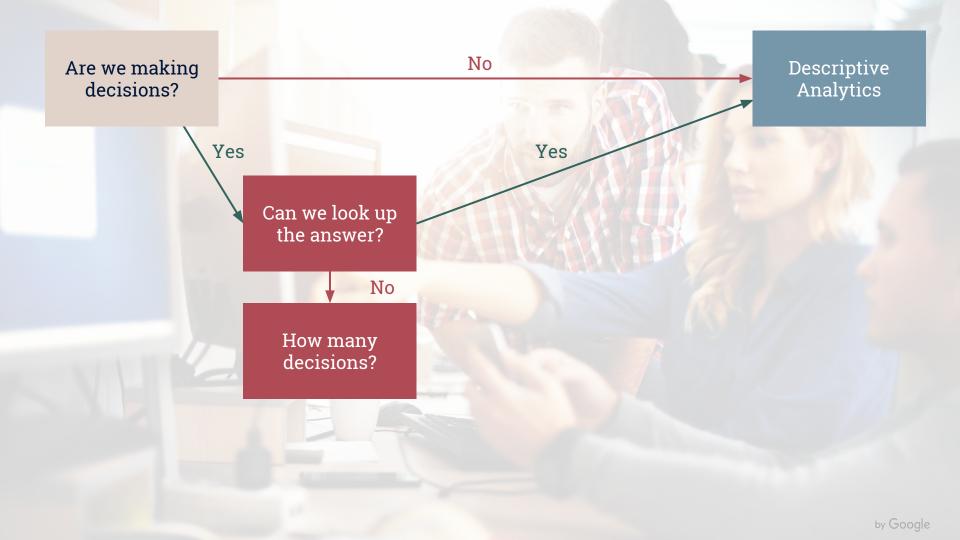


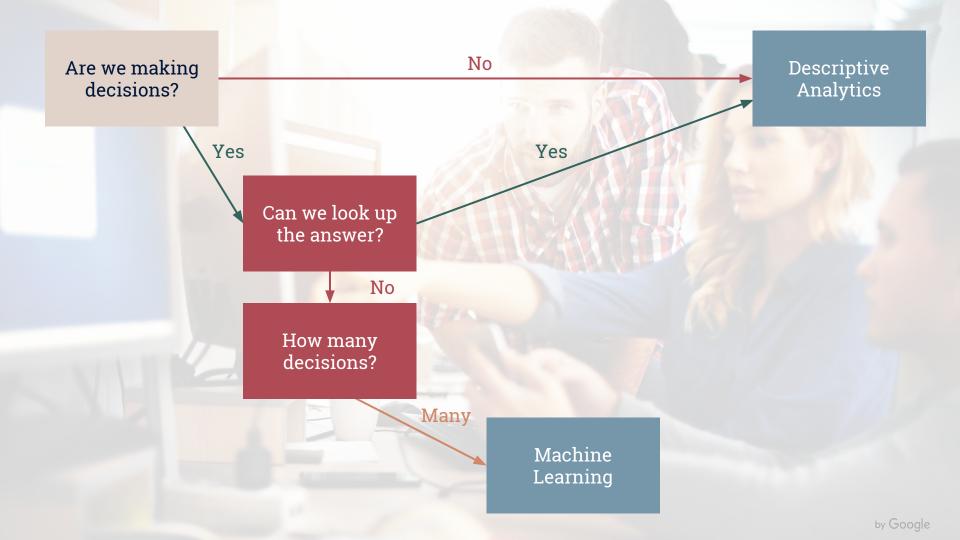


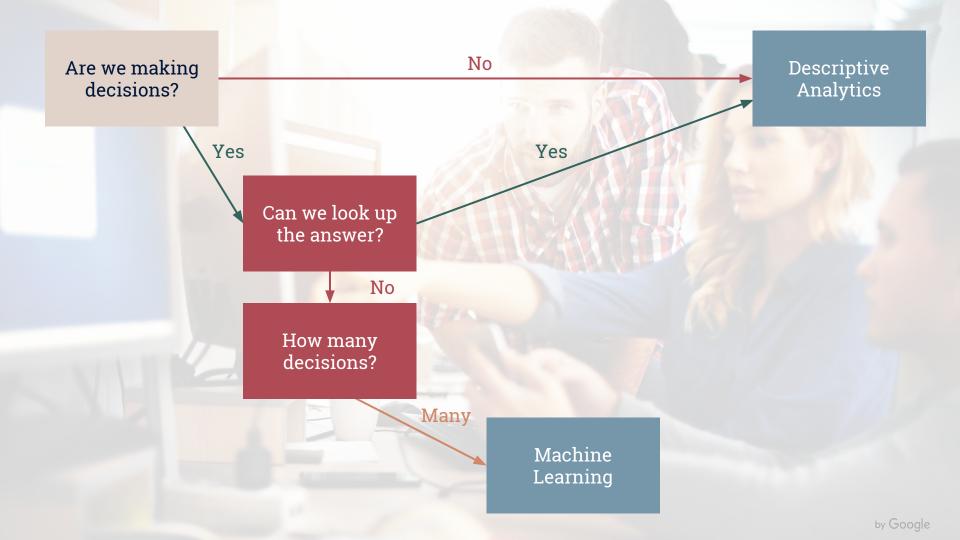


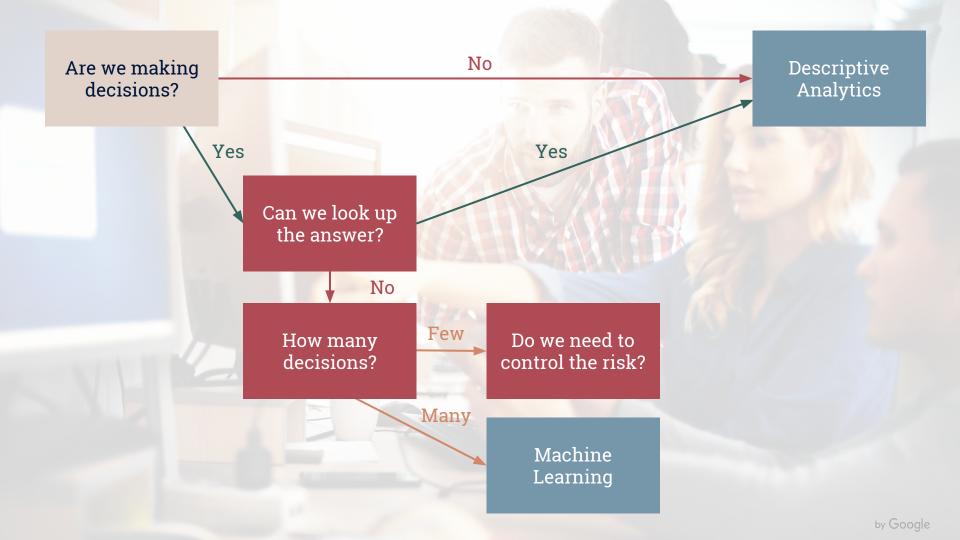


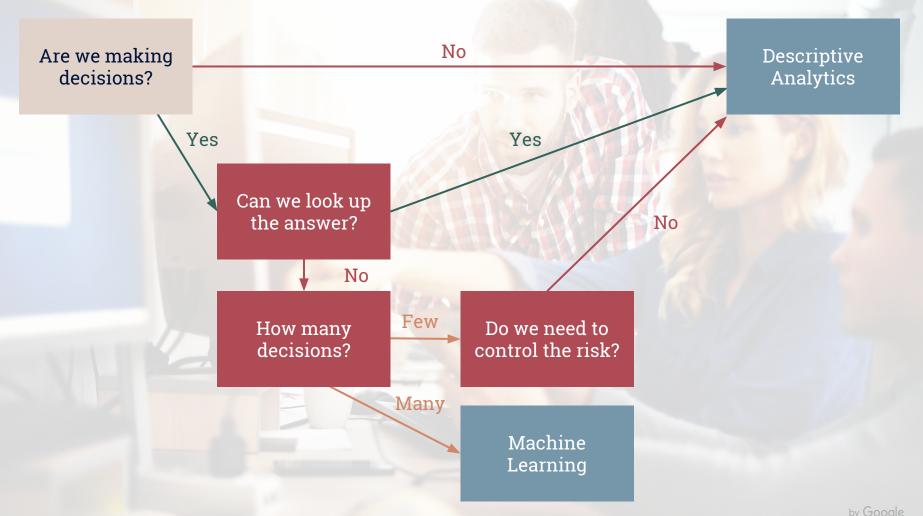


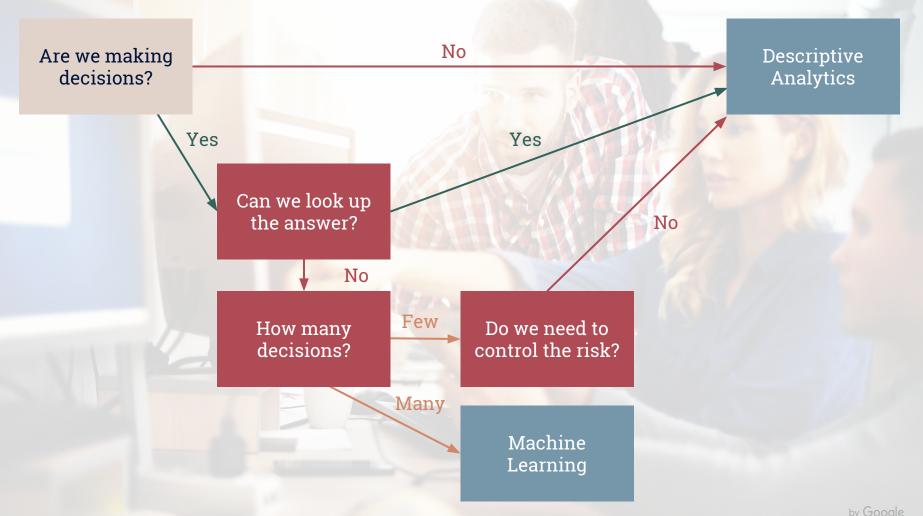


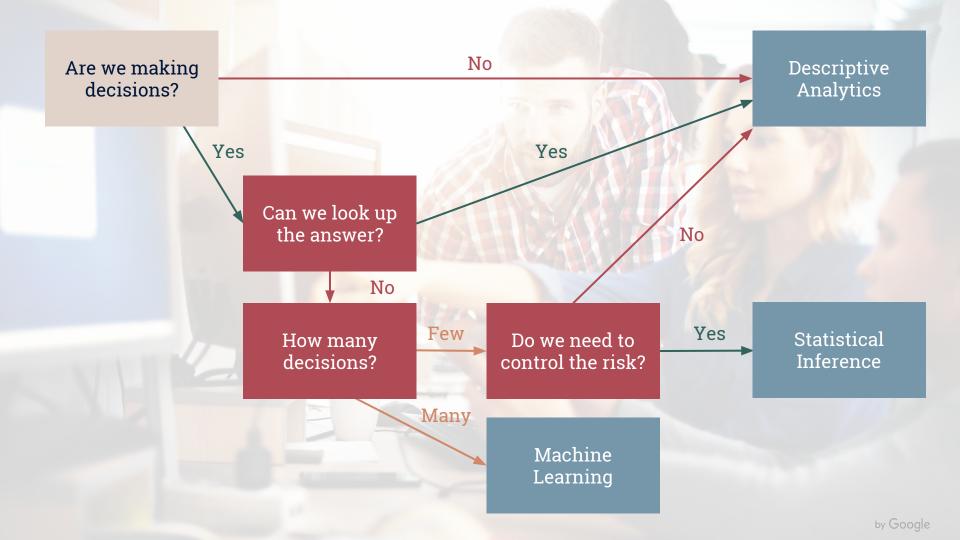












In Descriptive Analytics, there's only one mistake you can make: coming to conclusions.

If you're okay with sticking to what you know, you don't need a statistician.

"This is what's in my data.

All I know is that this is what's in my data."

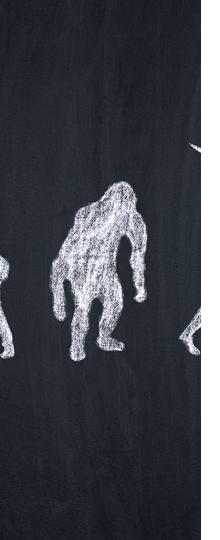


Key Ideas

Descriptive analytics is a good idea for all projects. Get started today!

Going beyond the data takes expertise.





Moving past Stage 2: No experts





Roles safari



#1 Decision-maker

Deep thinker "Choose actions sensibly."

#2 Data engineer

Excellent engineer "Get all the data."

#3 Basic analyst

Willing to learn "Plot some things."

#4 Expert analyst

Lightning fast "Plot all the things."

#5 Statistician

Rigorous & careful "Make conclusions safely."

#6 AI / ML engineer

Failure-tolerant "Make it work."

#7 Data scientist

Three-in-one "Explore, build, assess."

#8 Data science manager

"More impactful data science."

@quaesita

#9 Qualitative expert

Precise thinker "Clarify the decision."

@quaesita

#10 Researcher

Cutting edge "Design new algorithms."

Key Ideas

Experience and attitude are much more important than education or formal training.

A decision-maker who understands how to work with the other roles is an organization's best investment.





Moving past Stage 3: Isolated experts





Warning!

Avoid rigor for rigor's sake.



Rigor should start with the decision-maker.

Only go beyond descriptive analytics if a decision-maker has identified the need.

Inspiration is cheap and often good enough.





Get help!

From a teammate. From data.

Trouble getting started?

Hire a helper:

Metrics specialist

Hire a hybrid decision-maker:

Analytics manager Decision intelligence engineer



Help from the social sciences

1. Psychology

- a. Judgment & decision-making
- b. Social psychology
- c. Neuroeconomics

2. Economics

- a. Experimental game theory
- b. Behavioral economics
- 3. Managerial sciences

Role: Qualitative expert





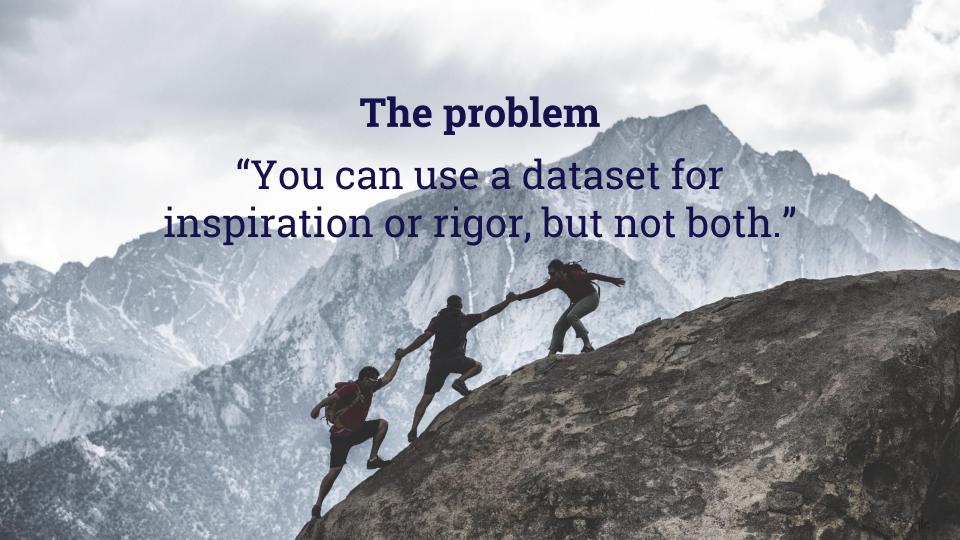
Get help!

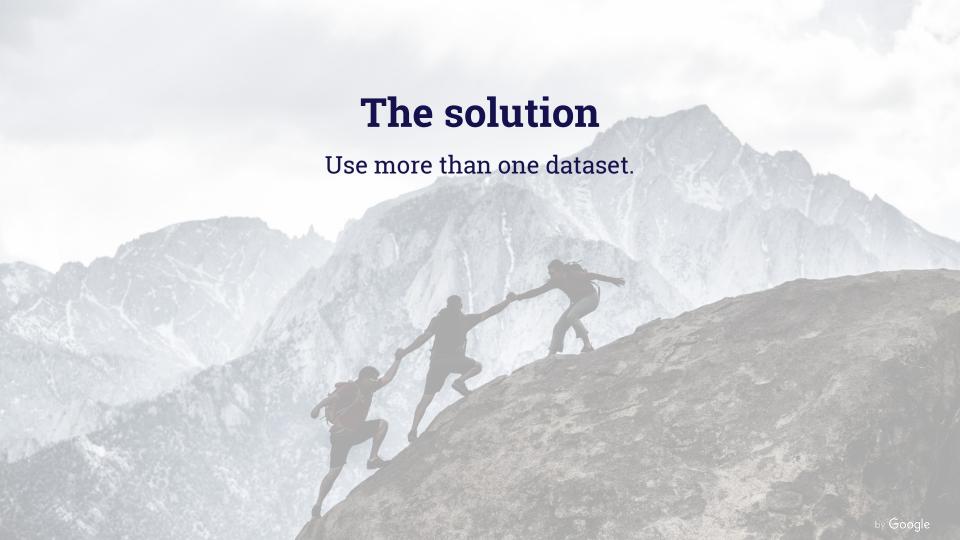
From a teammate. From data.

Warning!

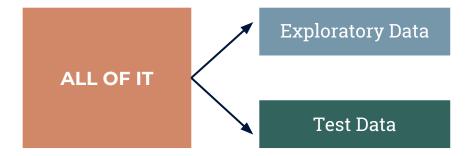
Data exploration can ruin **rigorous decision-making**.





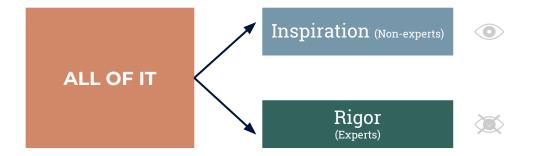


Always split your data before you begin.

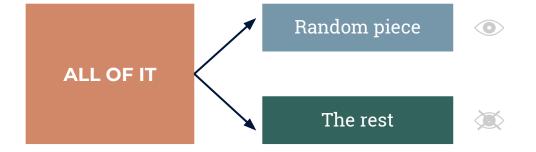




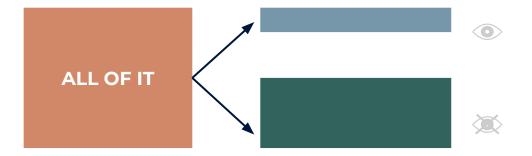
Always split your data before you begin.



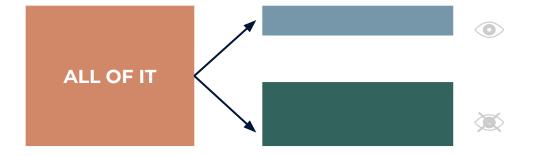




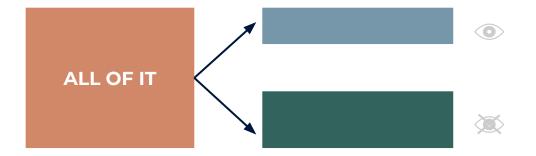




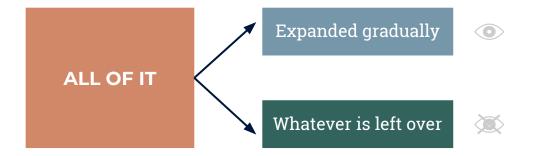




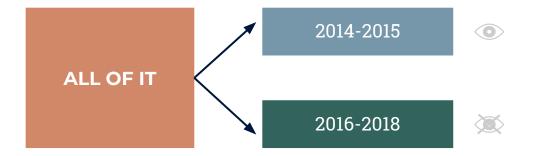




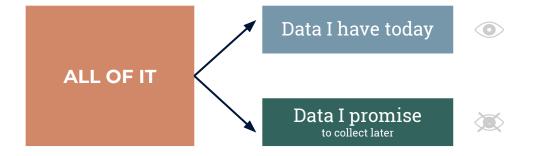








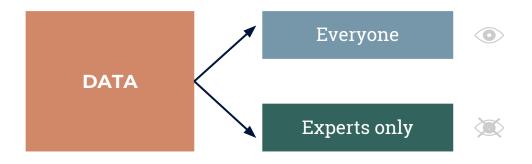






Key Ideas

Rigor starts with the decision-maker. Splitting your data is a cheap way to have your cake and eat it too.







Moving past Stage 4: Data nonsense



The solution

Appropriate delegation of decision responsibility.

Data-inspired versus data-driven.

4.2 ★★★★★ 5-star hotel

Hip upmarket hotel with pool & dining

Would you stay at this hotel?

4.2 ****

5-star hotel

Hip upmarket hotel with pool & dining

Data-inspired or data-driven?

Data-driven decision setup

No new information

Full information



Partial information



Data-driven decision setup

Default action



Full information



Partial information



No default action? Stick to descriptive analytics and go with your best guess.

Statistical inference is all about seeing if the evidence convinces us to change our minds.



Data-driven decision setup

Default action



Full information



Partial information



Data-driven decision setup

Default action

1

Metric design and criteria



Partial information



Create your metric before you name it.

Don't try to measure the platonic form of "happiness." Distill what's essential to your decision and call that metric whatever you like.



Data-driven decision setup

Default action

1

Metric design and criteria



Statistical considerations



Prevent nasty surprises by thinking about where your sample came from and how well it represents your population of interest.



Key Ideas

Design your decision process in this order:

Decision-making under no new information

→ Pick a default action

Decision-making under full information

→ Set decision criteria

Decision-making under partial information

→ Set statistical requirements

Too difficult? Stick to descriptive analytics.





Welcome to Stage 5: Data-driven!

Thank you!

Summary

- → Everyone is qualified to look at data. Say yes to descriptive analytics!
- → Default to a culture of data openness. Share where possible.
- → Inspiration is cheap, rigor is expensive.
- → Avoid jumping to conclusions when doing descriptive analytics.
- → Don't waste time on rigor unless the decision-maker is involved at the start.
- \rightarrow A habit of data splitting is the simplest and best cultural quick fix there is.
- → The only rule for splitting is: don't look at your test data.
- → No default action? Stick to descriptive analytics.
- → Create your metric before you name it.
- → Think about whether your sample adequately represents your population.

@quaesita by Google