



Open the Black Box: An Executive Guide to Making Unstructured Data Work in Finance



MICHELLE BONAT
CEO | CO-FOUNDER, DATA SIMPLY

@MBONAT

Agenda:

- Why
- What
- How

The Why:

A close-up portrait of Albert Einstein, showing his characteristic wild white hair and mustache. He is looking slightly to the right with a thoughtful expression, and his hands are clasped together in front of him. The background is a dark, textured grey.

Albert
Einstein

“The significant **problems we face** cannot be solved **at the same level of thinking** we are at when **we created them.**”

The Opportunity: How are these the same?



SEC Filings

8-K 1 twtr-8k_20160516.htm 8-K

**UNITED STATES
SECURITIES AND EXCHANGE COMMISSION**
Washington, D.C. 20549

FORM 8-K
CURRENT REPORT
Pursuant to Section 13 or 15(d) of
The Securities Exchange Act of 1934
Date of Report (Date of earliest event reported)
May 15, 2016

Twitter, Inc.
(Exact name of registrant as specified in its charter)

001-36164
(Commission File Number)

20-8913779
(IRS Employer Identification No.)

Delaware
(State or other jurisdiction of incorporation)

1355 Market Street, Suite 900
San Francisco, California 94103
(Address of principal executive offices, including zip code)

(415) 222-9670
(Registrant's telephone number, including area code)

Not Applicable
(Former name or former address, if changed since last report)

Check the appropriate box below if the Form 8-K filing is intended to simultaneously satisfy the filing obligation of the registrant under any of the following provisions (see General Instruction A.2. below):

☐ Written communications pursuant to Rule 425 under the Securities Act (17 CFR 230.425)

☐ Soliciting material pursuant to Rule 144-12 under the Exchange Act (17 CFR 240.144-12)

☐ Pre-commencement communications pursuant to Rule 144-2(b) under the Exchange Act (17 CFR 240.144-2(b))

☐ Pre-commencement communications pursuant to Rule 13e-4(c) under the Exchange Act (17 CFR 240.13e-4(c))

Indicate by check mark if the Registrant is a well-known seasoned issuer, as defined in Rule 405 of the Securities Act. Yes ☒ No ☐

Indicate by check mark if the Registrant is not required to file reports pursuant to Section 13 or Section 15(d) of the Act. Yes ☐ No ☒

Indicate by check mark whether the Registrant (1) has filed all reports required to be filed by Section 13 or 15(d) of the Securities Exchange Act of 1934 during the preceding 12 months (or for such shorter period that the Registrant was required to file such reports), and (2) has been subject to such filing requirements for the past 90 days. Yes ☒ No ☐

Indicate by check mark whether the Registrant has submitted electronically and posted on its corporate Web site, if any, every Interactive Data File required to be submitted and posted pursuant to Rule 405 of Regulation S-T (§232.405 of this chapter) during the preceding 12 months (or for such shorter period that the registrant was required to submit and post such

Table of Contents

EX-99.1 2 q116exhibit991er

SAN FRANCISCO – May 15, 2016 – Twitter, Inc. (“Twitter”) reported solid growth in our first quarter of 2016, with revenue increasing 137% year-over-year. Our revenue was primarily driven by our advertising business, which grew 137% year-over-year. Our revenue was primarily driven by our advertising business, which grew 137% year-over-year. Our revenue was primarily driven by our advertising business, which grew 137% year-over-year.

Lending Club today also reported solid growth in our first quarter of 2016, with revenue increasing 137% year-over-year. Our revenue was primarily driven by our advertising business, which grew 137% year-over-year. Our revenue was primarily driven by our advertising business, which grew 137% year-over-year.

Scott Sanborn will continue to serve as Chairman of the Board. Mr. Sanborn will be supported by the Board of Directors. Mr. Sanborn will be supported by the Board of Directors. Mr. Sanborn will be supported by the Board of Directors.

“A key principle of the Club is to ensure that the financial impact of this review is minimized. During the review, we will continue to operate as normal. We will continue to operate as normal. We will continue to operate as normal.”

“As our first quarter results demonstrate, Lending Club is well-positioned to continue its success. We are confident in our ability to continue to grow our business. We are confident in our ability to continue to grow our business. We are confident in our ability to continue to grow our business.”

First Quarter 2016 Financial Results

(\$ in millions)

Origination	Operating Revenue	Adjusted EBITDA ⁽¹⁾

They all can use:

machine learning to surface patterns for
data-driven decisions

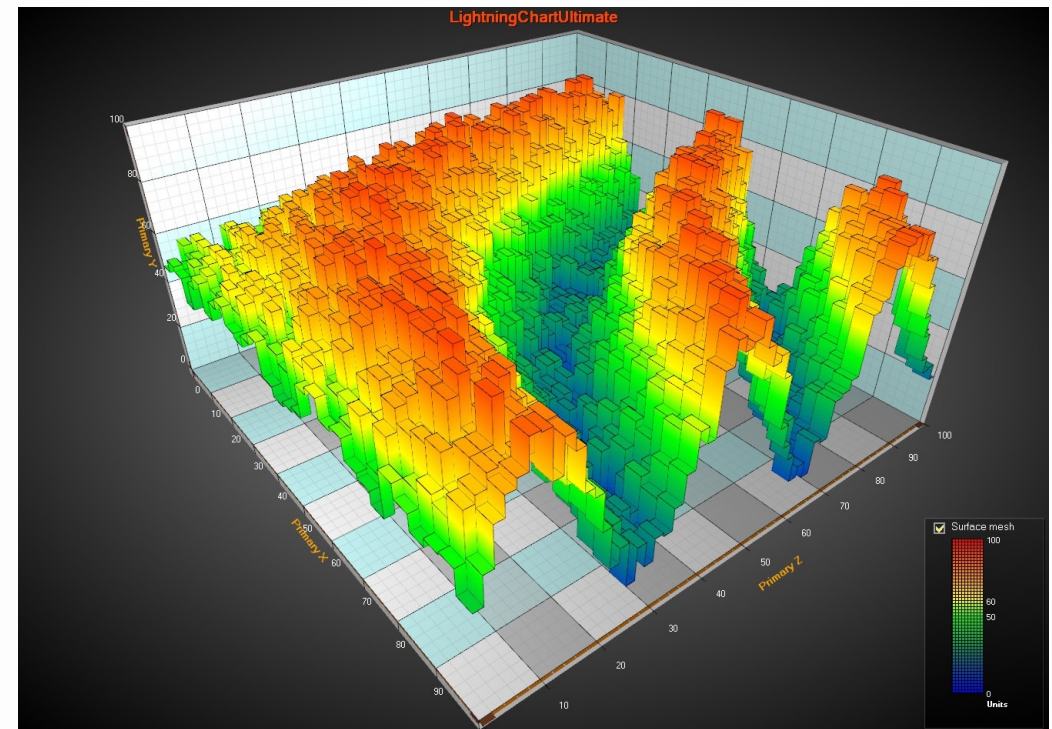
Which is better for making decisions?

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<feed xmlns="http://www.w3.org/2005/Atom">
<title>Latest Filings - Mon, 19 Sep 2016 13:29:19 EDT</title>
<link rel="alternate" href="/cgi-bin/browse-edgar?action=getcurrent"/>
<link rel="self" href="/cgi-bin/browse-edgar?action=getcurrent"/>
<id>http://www.sec.gov/cgi-bin/browse-edgar?action=getcurrent</id>
<author><name>Webmaster</name><email>webmaster@sec.gov</email></author>
<updated>2016-09-19T13:29:19-04:00</updated>
<entry>
<title>10-KT - Sino Fortune Holding Corp (0001628468) (Filer)</title>
<link rel="alternate" type="text/html"
href="http://www.sec.gov/Archives/edgar/data/1628468/000149315216013426/0001493152-16-013426-index.htm"/>
<summary type="html">
  <b>Filed:</b> 2016-09-19 <b>AccNo:</b> 0001493152-16-013426
  <b>Size:</b> 1 MB
</summary>
<updated>2016-09-19T13:16:08-04:00</updated>
<category scheme="http://www.sec.gov/" label="form type" term="10-K" />
<id>urn:tag:sec.gov,2008:accession-number=0001493152-16-013426</id>
</entry>
<entry>
<title>10-K/A - AMERICAN SOIL TECHNOLOGIES INC (0001031896) (Filer)
<link rel="alternate" type="text/html"
href="http://www.sec.gov/Archives/edgar/data/1031896/00011655271600000881-index.htm"/>
<summary type="html">
  <b>Filed:</b> 2016-09-19 <b>AccNo:</b> 0001165527-16-000881
  <b>Size:</b> 26 KB
</summary>
<updated>2016-09-19T13:00:53-04:00</updated>
<category scheme="http://www.sec.gov/" label="form type" term="10-K" />
<id>urn:tag:sec.gov,2008:accession-number=0001165527-16-000881</id>
</entry>
<entry>
<title>10-K - Crown Equity Holdings, Inc. (0001103833) (Filer)</title>
<link rel="alternate" type="text/html"
href="http://www.sec.gov/Archives/edgar/data/1103833/00014779321601012552-index.htm"/>
<summary type="html">
  <b>Filed:</b> 2016-09-19 <b>AccNo:</b> 0001477932-16-012552
  <b>Size:</b> 1 MB
</summary>
<updated>2016-09-19T06:27:22-04:00</updated>
<category scheme="http://www.sec.gov/" label="form type" term="10-K" />
<id>urn:tag:sec.gov,2008:accession-number=0001477932-16-012552</id>
</entry>
<entry>
<title>10-K/A - Mexus Gold US (0001355677) (Filer)</title>
<link rel="alternate" type="text/html"
href="http://www.sec.gov/Archives/edgar/data/1355677/000107878216003487/0001078782-16-003487-index.htm"/>
<summary type="html">
  <b>Filed:</b> 2016-09-19 <b>AccNo:</b> 0001078782-16-003487
  <b>Size:</b> 6 MB
</summary>
</entry>
```

<- This

OR

This ->



Of course!

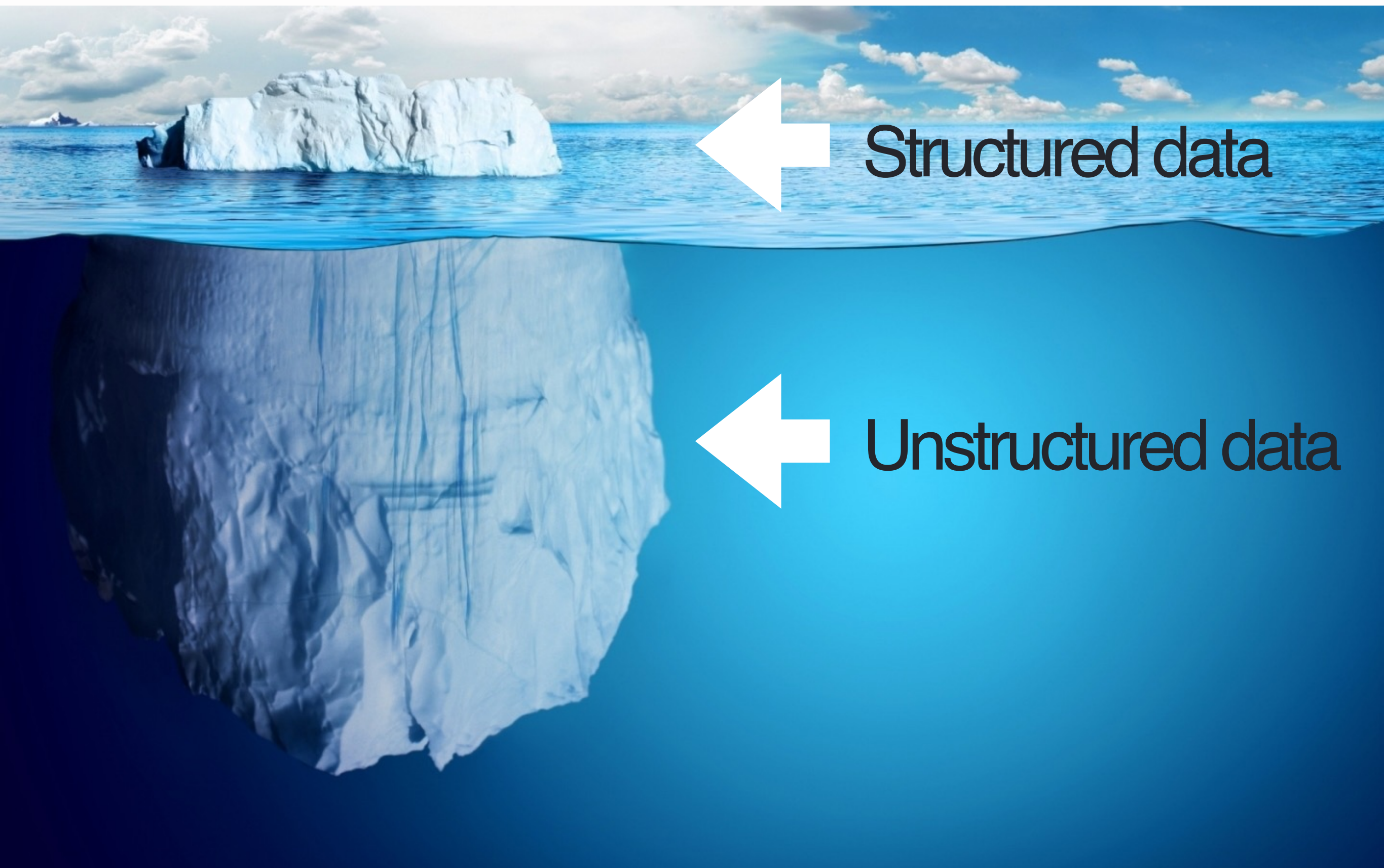
But how we do we get to data-driven
finance and risk with unstructured data?

The Challenge:



“My profession has probably been transformed again just since we started this session.”

Vast majority of the world's data is unstructured



Structured data

Unstructured data



WIKIPEDIA
The Free Encyclopedia

[Main page](#)
[Contents](#)
[Featured content](#)
[Current events](#)
[Random article](#)
[Donate to Wikipedia](#)
[Wikipedia store](#)

[Article](#)

[Talk](#)

[Read](#)

[Edit](#)

[View history](#)



Unstructured data

From Wikipedia, the free encyclopedia

Unstructured data (or **unstructured information**) refers to information that either does not have a pre-defined **data model** or is not organized in a pre-defined manner. Unstructured information is typically **text-heavy**, but may contain data such as dates, numbers, and facts as well. This results in irregularities and **ambiguities** that make it difficult to understand using traditional programs as compared to data stored in fielded form in databases or **annotated** (**semantically tagged**) in documents.

3 real world examples

- Save time
- Save money
- Get sued less

many more possibilities

Use Case #1: More Efficient Equity Research

- Track companies or tickers for bad and good signals.
- Know what to focus on.
- See trends and patterns for companies and sectors.
- **Benefits: Increase efficiency, be proactive.**

Use Case #2: Better Credit Risk Management

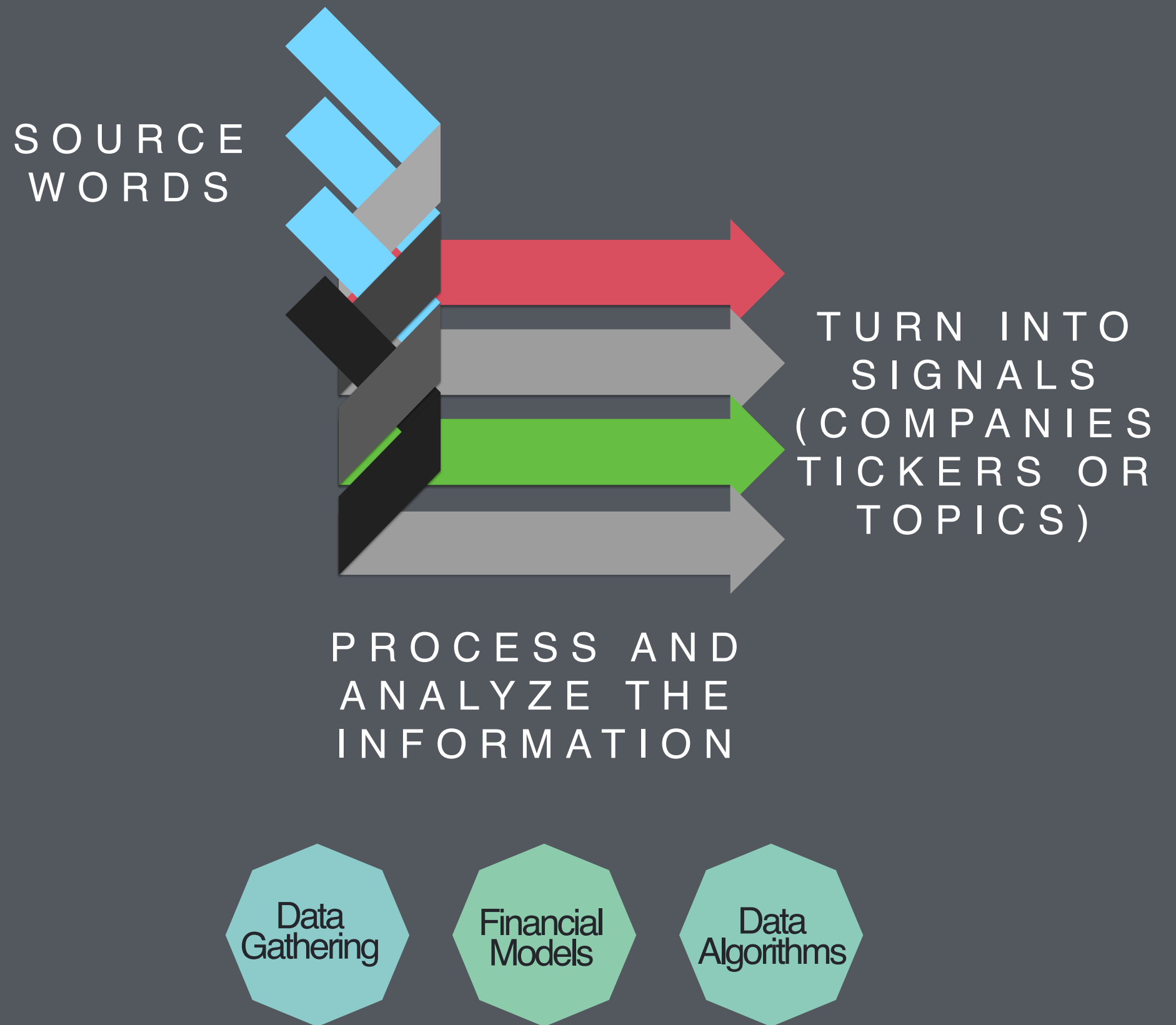
- More easily monitor issuers of holdings to see bad and good occurrences.
- Reduce downgrade surprises, be ahead of the curve.
- See negative trends earlier with history and comparison to peers.
- **Benefits: Reduce operational risk, increase efficiency, be proactive.**

Use Case #3: Proactive Mining of internal documents for risk and compliance (email, loan agreements, etc.)

- Proactively identify issues.
- Respond better to legal issues.
- Reduce manual human review of paper documents by external attorneys.
- See trends and patterns to catch issues earlier.
- **Benefits: Potential cost savings: \$\$\$ Millions of dollars, get sued less.**

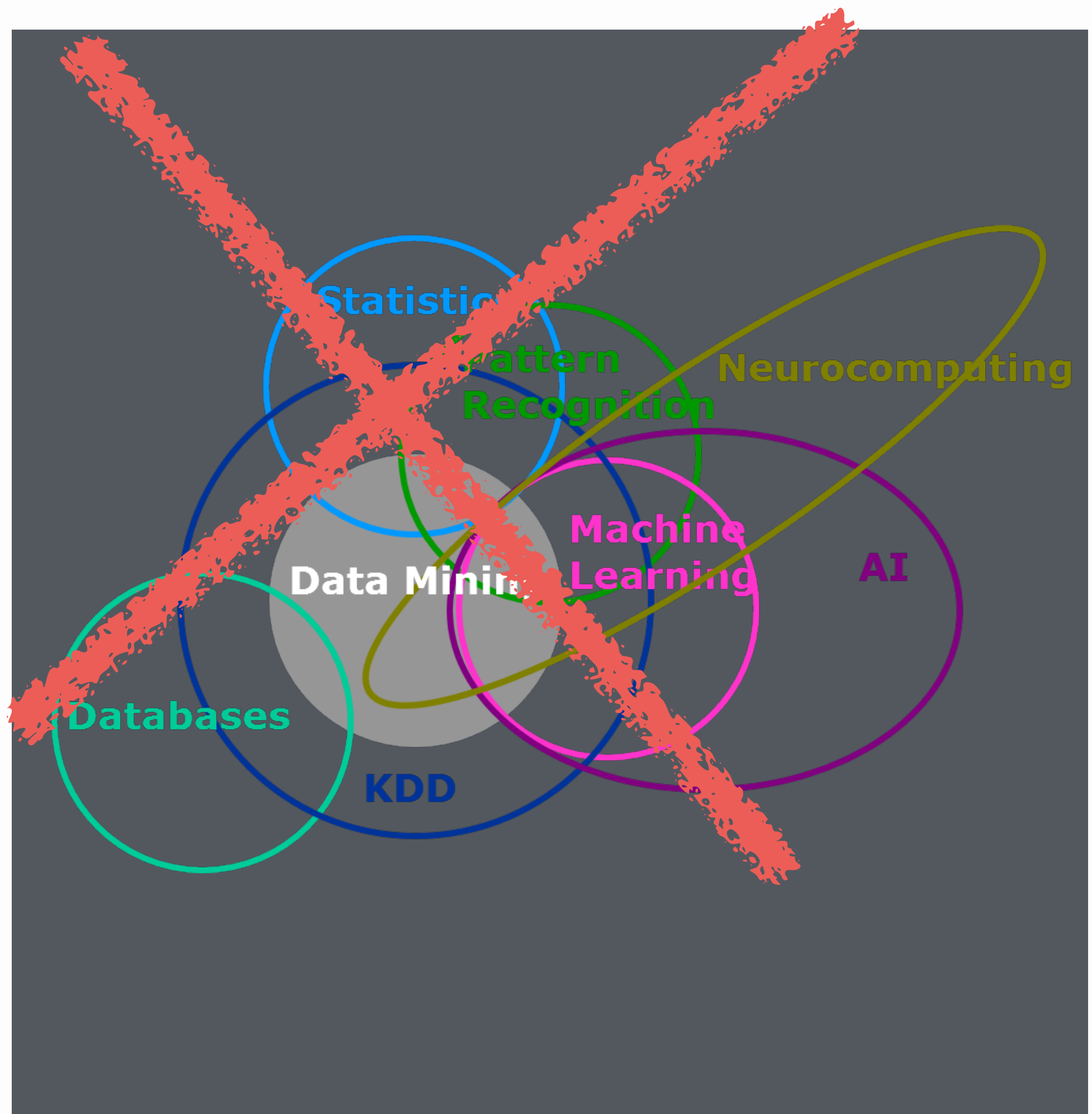
And many more possibilities ...

How The Tech Works



Step 1:

Ignore
Jargon



Step 2:

Be
Strategic:

Decide to
Ride the
Jet, not
Build It



Step 3:

Focus on Business Problem or Opportunity

- Cost efficiencies
- Drive revenue
- Compliance
- Risk management
- Innovation



Step 4:

Ask Key Questions



- Can somebody mine data for me and provide as a service, via signals?
- Can I easily consume the signals (API, app, or both)?
- Does it have built in understanding of financial terminology?
- Does it have comprehensive coverage for the data set I am evaluating?
- Can I find an engine to surface patterns in my own data?
- Can I easily mix different data sources together?

Summary:

- Leverage unstructured data in finance because it unlocks opportunities and solves problems in new ways.
- Unstructured data represents most of the data available but it is hard to make it usable.
- You need the data in the form of signals and patterns.
- You can use public data or mine your own.
- Many use cases: Equity analysis, credit risk, proactive compliance, and many more.
- Follow our steps to get there 1) Ignore jargon 2) Be strategic ride the jet don't build it 3) Focus on biz problem or opportunity 4) Ask key questions.

**GET THESE SLIDES:
INFO@DATASIMPLY.COM**

**TWITTER: @MBONAT @DATASIMPLY
WWW.DATASIMPLY.COM**