



Hands-On Introduction to Data-Science in Investment Management



CFA Society of Columbus
October 27th 2020



Tal Sansani, CFA

Data-scientist, quantitative analyst, and investment management consultant.

tsansani@gmail.com
<https://github.com/talsan>

- **UCLA, 2001-2005**
 - Applied Mathematics, Specialization in Computing
 - **American Century Investments, 2005-2019**
 - VP & Senior Quantitative Analyst
 - Quantitative Equity Group: Managing ~ \$12B AUM, across a range of Long-Short, Long-Only, Domestic, International products
 - Built & Managed the Data Science & Analytics Team
 - Research Focus: Sentiment signals, economic networks, trend-detection, intangible assets, and other non-financial indicators
 - **Investment Management Consultant, Present**
 - Help investment teams get more out of data, analytics, and technology.
-

Contents

1. **Practical Perspectives** (10 minutes)
2. **Exploratory Analysis of Conference Call Sentiment Models** (35 minutes)
3. **Q & A** (15 minutes)

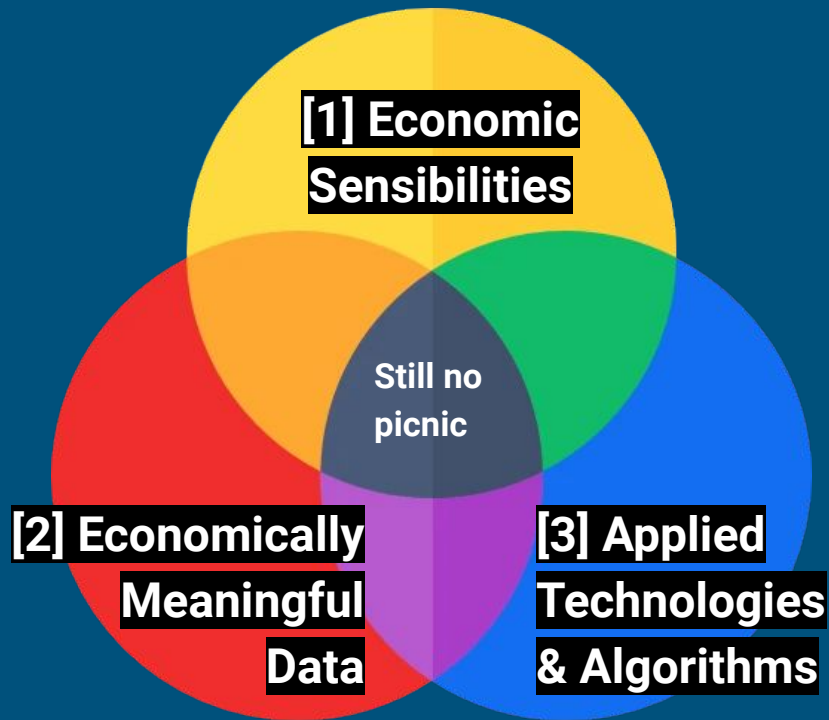
Practical Perspectives on Investment Data-Science

(but mostly just personal opinions)







Start with the Investment Problem First

Data-Science & Technology are a Means not an End

- Without economic sensibilities there is no anchor for understanding performance, incorporating feedback, and communicating with clients.
- Overfitting is the enemy
- A Data-Set is not a stock-selection idea
 - Like any other piece of information: it's how about how you contextualize it into your unique process



My Evolving Quant Research Stack

Big Data Storage, Processing & Power Tools:	AWS (S3, Athena, RDS, DynamoDB, EC2, Lambda, ElasticSearch)... with nothing against Google or Azure, but as you do more, it becomes more cost-effective to pick one.	
Data Engineering, Web Interfacing & Modelling:	Python (but that doesn't mean R can't handle it, it's just a bit more work to make it play nice with other systems).	
Interactive Development Environment (IDE):	Pycharm for python data/model work; rStudio for everything R	
Analytics & Data-Exploration:	RStudio is great (<u>tidyverse</u> makes analysis easy), but I still use python because I don't want to manage two environments.	
Reproducible Research Reports:	Jupyter Notebook & R Markdown (note: these are amazing tools, but not development environments!!!)	
Interactive Dashboards:	Tableau for traditional dashboards & canned reports (powerful charts in no-time); Shiny/Dash/Flask for highly customized research applications (endlessly flexible);	

How to Get Started?

Little Steps That Add Up

1. **Champion a data-driven culture** and awareness of modern investment applications of data & technology
 - Meaningfully connect data technologists with investment management
2. **Start with a small, targeted, actionable investment idea** or business-process-improvement
 - Avoid big splashes, mass hirings, and giant meetings
3. **Evaluate the value**, the costs, and learn from the pain points
4. **Repeat Steps 1 through 3**
 - Adding resources (talent, data, and technology) along the way

Hands-On Introduction Textual-Analysis w/ Conference Calls

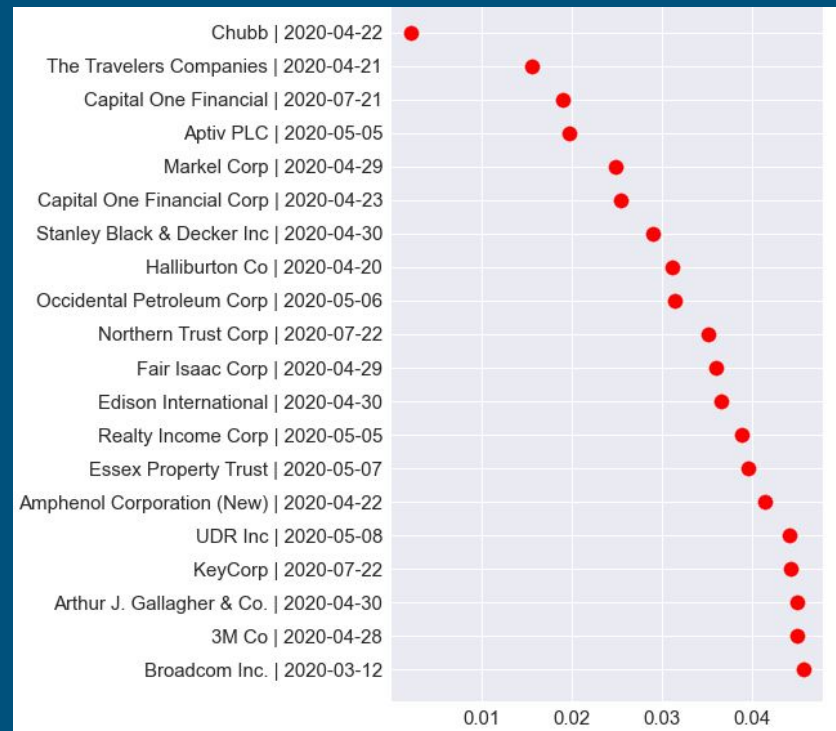
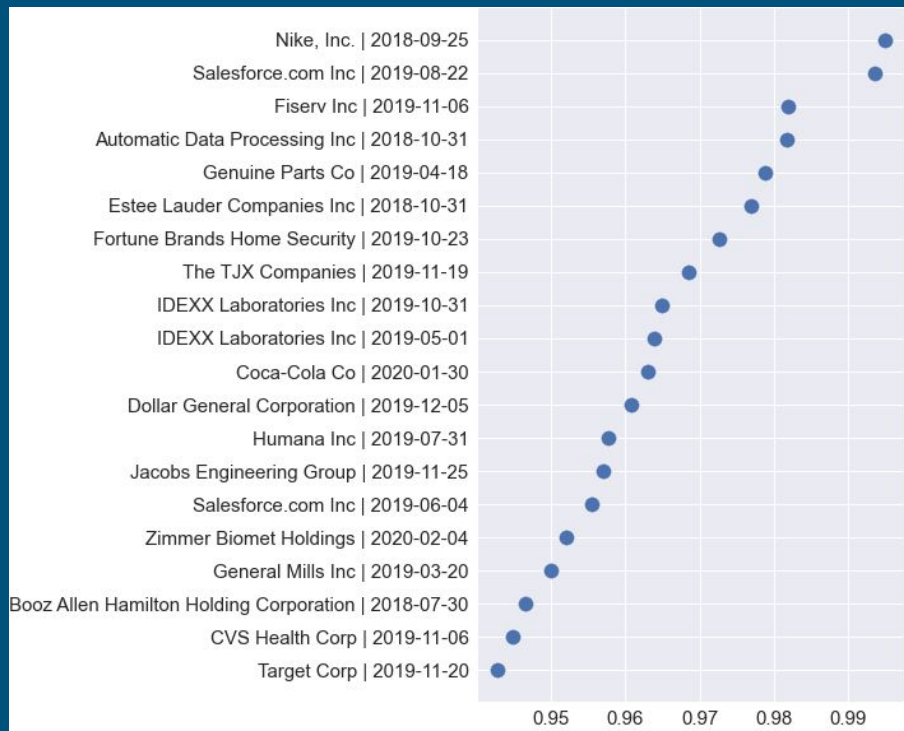
Part I: Data Sourcing & Structuring

Part II: Three Sentiment Models

Part III: Exploratory Analysis

<https://github.com/talsan/calltone>

Top / Bottom Sentiment Conference Calls Over the Last Two Years



... so how did we get here?

Part II: Quant Research Data Flow

Data Sourcing & Structuring



The Motley Fool

<https://github.com/talsan/foolcalls>

iShares

by BLACKROCK®

<https://github.com/talsan/ishares>



**ALPHA
VANTAGE**

<https://github.com/talsan/alphavantage>

AWS S3 Data Lake

Transcripts,
metadata, speaker
info, etc.

universe, sectors,
style, size, cusips,
sedols, etc.

split/dividend
adjusted asset
returns



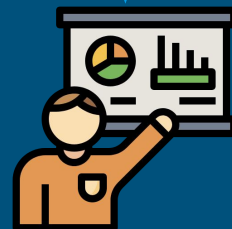
**AWS Athena
Analytics (SQL-Like)**



**Conference Call
Sentiment & Analysis**



<https://github.com/talsan/calltone>



Appendix

Appendix A: Building & Cultivating an Investment Data-Science Team

Hiring Talent

- Find candidates who show genuine interest in the investing domain
- Set realistic expectations
- Don't get too hung up in specific programming languages or technologies

Cultivating Talent

- Tightly integrate Data scientists and technologists with their investment peers
- Align incentives and blur the multi-disciplinary lines because **it's all connected** -- data, research, portfolio management, analytics -- and serving a common goal

Appendix B: Applications of Conference Calls

Investment Applications


1. **Tone:**
Reflects management's optimism (or pessimism) about future prospects
2. **Communication Quality / Manipulation:**
Are they exaggerating? Obfuscating?
Omitting key information?
3. **Linkages & Text-Based Peer Groups**
4. **Emerging Themes / Topic Analysis**

Contextual Integration

- **How might this help a quant?**
 - Adds much needed context/“color” to a framework that (usually) invests on numeric/trailing information.
 - Ability to scale systematic insights
- **How might this help a fundamental team?**
 - Bubble up “needle in the haystack” red-flags or insights (especially among less covered names)
 - Streamline research process
 - Provide a “2nd opinion” and potentially reveal biases

Appendix C: Fool.com Structuring Example

KRUS earnings call for the period ending May 31, 2020.

 Motley Fool Transcribers (MFTranscribers)
00 Jul '20 at 1:35AM

Kura Sushi USA Inc ([NASDAQ:KRUS](#))
Q3 2020 Earnings Call
Jul 14, 2020, 4:30 p.m. ET

Contents:

- Prepared Remarks
- Questions and Answers
- Call Participants

Prepared Remarks:
Operator

Good day, ladies and gentlemen, and thank you for standing by. Welcome to the Kura Sushi USA, Inc. Third Quarter 2020 Earnings Conference Call. At this time, all participants have been placed in a listen-only mode. Please note that this call will be recorded. If you have any questions, please hold on until the question-and-answer session. Thank you for your participation, and we will now begin the call. Please note that this call is for informational purposes only and should not be relied upon for investment decisions. Please note that this call is for informational purposes only and should not be relied upon for investment decisions.

On the call today we have Jimmy Uba, President and Chief Executive Officer; Officer; and Benjamin Porten, Investor Relations Manager.

I would now like to turn the conference over to Mr. Porten. Thank you. You may begin.

Benjamin Porten -- Investor Relations Manager

Thank you, operator. Good afternoon everyone and thank you all for joining. access to our fiscal third quarter 2020 earnings release. It can be found at our Investor Relations section. A copy of the earnings release has also been included in our earnings presentation. Thank you, and we will now begin the call.

```
{
  "publication_author": "Motley Fool Transcribers (MFTranscribers)",
  "publication_time_published": "2020-07-15 05:30:00",
  "publication_time_updated": "2020-07-15 05:30:00",
  "call_title": "Kura Sushi USA Inc (KRUS) Q3 2020 Earnings Call Transcript",
  "call_subtitle": "KRUS earnings call for the period ending May 31, 2020.",
  "period_end": "2020-05-31",
  "ticker": "KRUS",
  "ticker_exchange": "NASDAQ",
  "company_name": "Kura Sushi USA Inc",
  "fool_company_id": "341476",
  "fiscal_period_year": "2020",
  "fiscal_period_qtr": "Q3",
  "call_short_title": "Q3 2020 Earnings Call",
  "call_date": "2020-07-14",
  "call_time": "2020-07-14 16:30:00",
  "duration_minutes": "47",

```

```
"call_transcript": [
  {
    "statement_num": 1,
    "section": "pres",
    "statement_type": "O",
    "speaker": "Operator",
    "role": "",
    "affiliation": "",
    "text": "Good day, ladies and gentlemen, and thank you for standing by."
  },
  {
    "statement_num": 2,
    "section": "pres",
    "statement_type": "P",
    "speaker": "Benjamin Porten",
    "role": "Investor Relations Manager",
    "affiliation": "",
    "text": "Thank you, operator. Good afternoon everyone and thank you all for joining."
  }
]
```

Call-Level Meta-Data

Transcripts
(with speaker and
statement metadata)

Appendix D: AWS Athena Screen Shot

The screenshot displays the AWS Athena Query Editor interface. The top navigation bar includes the AWS logo, 'Services', and a dropdown menu. The main header shows 'Athena' and 'Query editor' with tabs for 'Saved queries', 'History', 'Data sources', and 'Workgroup : qc'. On the right, there are links for 'Settings', 'Tutorial', 'Help', and 'What's new', along with a '10+' badge.

The left sidebar contains a 'Data source' section with a dropdown menu set to 'AwsDataCatalog' and a 'Connect data source' link. Below this is a 'Database' section with a dropdown menu set to 'qcdb' and a 'Filter tables and views...' input field. The 'Tables (9)' section lists various tables, including 'av_invalid_tickers', 'av_prices', 'edgaraws_filing_metadata', 'edgaraws_masteridx', 'etf_holdings', 'fool_call_index', 'fool_call_speakers_nested', 'fool_call_statements_nested', and 'ishares_holdings'. The 'Views (4)' section lists 'av_pricing', 'edgarfilinginfo_v', 'fool_call_speakers', and 'fool_call_statements'.

The main query editor area shows a SQL query: `1 | SELECT * FROM "qcdb"."fool_call_statements" limit 10;`. Below the query, there are buttons for 'Run query', 'Save as', 'Create', 'Format query', and 'Clear'. The 'Run query' button is highlighted. The status bar indicates 'Run time: 1.93 seconds, Data scanned: 16.96 MB'.

The 'Results' section displays a table with the following columns: 'statement_num', 'section', 'statement_type', 'speaker', and 'role'. The table contains 10 rows of data, showing details for various statements from the '19-mercantile-bank-mbwm-q2-2018-earnings-conference-c'.

	statement_num	section	statement_type	speaker	role
19-mercantile-bank-mbwm-q2-2018-earnings-conference-c	1	pres	O	Operator	
19-mercantile-bank-mbwm-q2-2018-earnings-conference-c	2	pres	P	Mike Houston	Lambert, Edwards & Asso
19-mercantile-bank-mbwm-q2-2018-earnings-conference-c	3	pres	P	Robert B. Kaminski Jr.	President and Chief Execu
19-mercantile-bank-mbwm-q2-2018-earnings-conference-c	4	pres	P	Raymond E. Reitsma	Chief Financial Officer
19-mercantile-bank-mbwm-q2-2018-earnings-conference-c	5	pres	P	Charles E. Christmas	Chief Financial Officer
19-mercantile-bank-mbwm-q2-2018-earnings-conference-c	6	pres	P	Robert B. Kaminski Jr.	President and Chief Execu
19-mercantile-bank-mbwm-q2-2018-earnings-conference-c	7	qa	O	Operator	
19-mercantile-bank-mbwm-q2-2018-earnings-conference-c	8	qa	Q	Brendan Nosal	Analyst