Data Science Notes

Barclays Research's Data & Investment Sciences blog, where we dive into some of the frequently asked questions about data science and how it is reshaping investing.

BARCLAYS

29 June 2020

Introducing Data Science Notes

Investment Sciences
Ben McSkelly, CFA

Ryan Preclaw

Data Science
Adam Kelleher

Welcome to Data Science Notes, Barclays Research's Data and Investment Sciences blog that surveys topics of interest to investors looking to leverage alternative data and advanced empirical methods to inform their investment process.

Why read Data Science Notes?

Data Science Notes draws on the experience of Barclays' Data and Investment Sciences teams, with insights on how to build a data science capability as well as implement a data-led investment strategy. There are numerous unique opportunities and challenges presented by using alternative data in our industry. However, Data Science Notes does not provide investment insights, but instead considers the opportunities and challenges involved in using alternative data to generate next-level insights and actionable ideas. Topics include:

- **Infrastructure:** We discuss how to choose the most appropriate infrastructure to store and analyse big data in a compliant manner.
- **Data sources:** We discuss the alternative data types that exist in the market and what benefits they bring, as well as highlight in more detail specific data types.
- **Statistical and data methods:** We dive into the statistical and analytic methods used to process and analyse large datasets, such as machine learning, NLP, and classification models.
- **Operational issues:** Using alternative data in a regulated investment business can present a number of unique legal, compliance and general operational issues such as building a suitable team, preserving data privacy and vetting vendors.

What is Barclays Research Data and Investment Sciences?

Leveraging alternative and big data, the Data and Investment Sciences teams are embedded within Barclays Investment Bank Research, providing institutional clients with insights and actionable ideas.

Two teams working side by side: a pure Data Science team and an Investment Sciences team looking to integrate data science into the investment process

These two teams of 10+ data and investment scientists, based in New York and London, work in parallel. The Data Science team operates the dynamic infrastructure, processes large datasets and builds analytical and statistical methodologies. The Investment Sciences team focuses on using these new techniques to answer critical investment research questions and develop and validate investment theses.

Together with our fundamental analysts, we provide clients written research with security-specific analysis, recommendations and thematic insights.

Additionally, we share our methodologies. Transparency is key to what we do.

The four pillars of our blog

Our Data Science Notes will cover four areas of interest to those looking to implement a data-led strategy in investing or investment research:

Infrastructure

This pillar focuses on how to scope out, implement and operate the infrastructure required for big data in a compliant fashion. A one-size-fits-all solution does not exist; instead, a variety of options for data storage, computing and services are available for different scales and requirements. Topics include where cloud solutions are appropriate, how to implement agile development principles in a research setting as well as the benefits of a microservice architecture.

Data sources

We see the world of investment data in five broad categories. First, the foundation is traditional fundamental data, including pricing, returns, accounting data, economics series and estimates. Second, third, and fourth are the large alternative data types that can potentially be applied across a large proportion of the market (but have significant costs of entry):

- Geolocation data
- Transaction data (commonly, payment card data)
- NLP on text sources (news, social media, corporate transcripts)

Fifth is the long tail of alternative data with narrower uses: industry or company-specific, open datasets provided by governments or self-created sets such as web scraping products.

In future blog posts we will discuss these topics in more detail, as well as our approaches to the challenges and opportunities presented by these datasets.

Statistical and data methods

With new unique datasets, often from outside the traditional fundamental sources, robust statistical techniques are key to obtaining reliable answers. The foundation is statistical best practises, particularly for time series data and forecasting. Then, an analysis can be expanded to include appropriate techniques such as classification models, various machine learning frameworks, and complex algorithm sets such as natural language processing (NLP).

Operational issues

There are specific legal, staffing and compliance considerations that arise from the use of alternative data and cloud computing, as well as the storage of data to ensure it does not contain personally identifiable information. We aim to cover these topics as well as how to hire the personnel to build a data science capability.

In future posts we will discuss these topics in more detail and our approaches to the challenges and opportunities presented by alternative data.

Disclaimer:

BARCLAYS

This communication has been prepared by Barclays.
"Barclays" means any entity within the Barclays Group of companies, where "Barclays Group" means Barclays Bank PLC, Barclays PLC and any of their subsidiaries, affiliates, ultimate holding company and any subsidiaries or affiliates of such holding company.

CONFLICTS OF INTEREST

BARCLAYS IS A FULL SERVICE INVESTMENT BANK. In the normal course of offering investment banking products and services to clients, Barclays may act in several capacities (including issuer, market maker and/or liquidity provider, underwriter, distributor, index sponsor, swap counterparty and calculation agent) simultaneously with respect to a product, giving rise to potential conflicts of interest which may impact the performance of a product.

NOT RESEARCH

The information provided does not constitute 'investment research' or a 'research report' and should not be relied on as such, although it may contain references to views or information published by the Barclays Research department. Investment decisions should not be based upon the information provided.

BARCLAYS POSITIONS

Barclays may at any time acquire, hold or dispose of long or short positions (including hedging and trading positions) and trade or otherwise effect transactions for their own account or the account of their customers in the products referred to herein which may impact the performance of a product.

FOR INFORMATION ONLY

This information has been prepared by the Research Department within the Investment Bank of Barclays. The information, analytic tools, and/or models referenced herein (and any reports or results derived from their use) are intended for informational purposes only. Barclays has no obligation to update this information and may cease provision of this information at any time and without notice.

NO OFFER

Barclays is not offering to sell or seeking offers to buy any product or enter into any transaction. Any offer or entry into any transaction requires Barclays' subsequent formal agreement which will be subject to internal approvals and execution of binding transaction documents. The products mentioned may not be eligible for sale in some states or countries, nor suitable for all types of investors.

NO LIABILITY

Neither Barclays nor any of its directors, officers, employees, representatives or agents, accepts any liability whatsoever for any direct, indirect or consequential losses (in contract, tort or otherwise) arising from the use of this communication or its contents or reliance on the information contained herein, except to the extent this would be prohibited by law or regulation.

NO ADVICE

Barclays is not acting as a fiduciary. Barclays does not provide, and has not provided, any investment advice or personal recommendation to you in relation to any transaction and/or any related securities described herein and is not responsible for providing or arranging for the provision of any general financial, strategic or specialist advice, including legal, regulatory, accounting, model auditing or taxation advice or services or any other services in relation to the transaction and/or any related securities described herein. Accordingly Barclays is under no obligation to, and shall not, determine the suitability for you of the transaction described herein. You must determine, on your own behalf or through independent professional advice, the merits, terms, conditions and risks of any transaction described herein.

NOT A BENCHMARK

The information provided does not constitute a financial benchmark and should not be used as a submission or contribution of input data for the purposes of determining a financial benchmark.

INFORMATION PROVIDED MAY NOT BE ACCURATE OR COMPLETE AND MAY BE SOURCED FROM THIRD PARTIES

All information is provided "as is" without warranty of any kind. Because of the possibility of human and mechanical errors as well as other factors, Barclays is not responsible for any errors or omissions in the information contained herein. Barclays is not responsible for information stated to be obtained or derived from third party sources or statistical services. Barclays makes no representation and disclaims all express, implied, and statutory warranties including warranties of accuracy, completeness, reliability, fitness for a particular purpose or merchantability of the information contained herein.

PAST & SIMULATED PAST PERFORMANCE

Any past or simulated past performance including back-testing, modelling or scenario analysis contained herein is no indication as to future performance. Past performance is not necessarily indicative of future results. No representation is made as to the accuracy of the assumptions made within, or completeness of, any modelling, scenario analysis or back-testing and no representation is made that any returns will be achieved.

OPINIONS SUBJECT TO CHANGE

All opinions and estimates are given as of the date hereof and are subject to change. The value of any investment may also fluctuate as a result of market changes. Barclays is not obliged to inform the recipients of this communication of any change to such opinions or estimates.

NOT FOR RETAIL

This document is being directed at persons who are professional investors and is not intended for retail customer use.
Not For Further Distribution or Distribution To Retail Investors.
For Discussion Purposes Only.

IMPORTANT DISCLOSURES

For important regional disclosures you must read, visit the link relevant to your region. Please contact your Barclays representative if you are unable to access.
EMEA
https://www.home.barclays/disclosures/important-emea-disclosures.html.
APAC
https://www.home.barclays/disclosures/important-apac-disclosures.html.
U.S.
https://www.home.barclays/disclosures/important-us-disclosures.html.

CONFIDENTIAL

This communication is confidential and is for the benefit and internal use of the recipient for the purpose of considering the securities/transaction described herein, and no part of it may be reproduced, distributed or transmitted without the prior written permission of Barclays.

ABOUT BARCLAYS

Barclays Bank PLC is authorised by the Prudential Regulation Authority and regulated by the Financial Conduct Authority and the Prudential Regulation Authority and is a member of the London Stock Exchange. Barclays Bank PLC is registered in England No. 1026167 with its registered office at 1 Churchill Place, London E14 5HP.

COPYRIGHT

© Copyright Barclays 2020 (all rights reserved).