Coin Metrics Data Science Interview Case Study

2019-11-11

Introduction

Thank you for your interest in Coin Metrics! We would like to invite you to take the following case study which is designed to provide you with an actual example of the work that Coin Metrics does while also giving us an opportunity to evaluate your skills and abilities. We hope this case study is interesting to you, gives you insight into the responsibilities of the position, while also not placing an unfair burden on your time.

Company Background

Coin Metrics was started in 2017 as a project aimed at providing the public with reproducible, clearly-defined cryptocurrency network metrics. The metrics we produce resemble traditional macroeconomic indicators that measure economic activity in nation states – in the same way that GDP, inflation rate, and unemployment rate provide insight into the vitality of a country's economy, transaction volume and active addresses allow one to assess the economic vibrancy of cryptocurrency networks, such as Bitcoin and Ethereum. Data produced by Coin Metrics attracted attention from large financial institutions and at the end of 2018, Coin Metrics formally incorporated as a company. Capital received from a seed round allowed us to dramatically expand the scope and depth of on-chain data collection and analysis, and establish a foundation of market data gathering, processing and shipment.

Despite the company's young age, we have signed contracts for delivery of network and market data with many financial institutions, notably, with multiple business units of Fidelity Investments. The company is led by CEO Tim Rice, who previously served as the Global Head of Pricing and Reference Data at Thomson Reuters.

Case Study Instructions

Coin Metrics publishes several network metrics and no single metric can fully represent activity on the network. The large volume of different metrics also makes it difficult to understand the current state of economic activity. Therefore, we are interested in creating an aggregate growth metric: a single, easily-interpretable network metric that comprehensively measures how fast a network is growing.

The ideal aggregate growth metric would use several of our network metrics as inputs and allow a user to determine whether a network is experiencing positive or negative growth as well as the magnitude of the growth. The metric would be comparable across time (allow a user to measure growth today versus some time in the past) and across cryptoassets (allow a user to compare the growth of Bitcoin versus Ethereum).

For inspiration, imagine a single growth indicator for a nation state that resembles real GDP growth or a purchasing managers' index. A country's business cycle can be clearly seen as growth regularly oscillates between periods of positive growth and negative growth.

The goal of the case study is to create an aggregate growth metric for Bitcoin. You can obtain a subset of our network data for Bitcoin using our Community Data, a free version of our product that we offer to the public. Commuinty Data can be downloaded in the form of a csv file using our API tools page. The tools page also allows you to view and download a description of each of the metrics we offer.

Here are some questions that you may encounter and some hints on how to address them.

1. Which network metrics should be used as inputs into the aggregate growth metric? For Bitcoin, growth was extremely strong at the end of 2017 but then turned negative for most of 2018. Some indicators like `TxCnt` are good candiates for inclusion while others like `VtyDayRet180d` are not.

2. Some network metrics like `TxCnt` exhibit a lot of noise and some seasonal patterns. What is the best way to lessen the impact of noise and seasonaity and amplify the signal in these metrics?
3. Many of our network metrics measure the level of something. What are the best set of transformations to transform a metric from level-space to growth-space?
4. Bitcoin's growth was much stronger earlier in its history because it started from a smaller base which makes visualization and comparison across time challenging. Is there a method of calculating growth that is suitable for large positive and negative growth numbers?

Please send code and results via a version-controlled repository or through email. If you have any questions, please do not hesitate to reach out. We very much are interested in reviewing your results.

Addendum

Coin Metrics understands that asking candidates to take programming assignments, while common practice in the industry, can place an unfair burden on candidates. If you have an portfolio or project that you feel demonstrates your skills and abilities well, please feel free to send that in leiu of taking the case study. We would also be willing to consider partial implementation of the case study with well-thought answers to the questions raised above.