Analytics mindset

EDGAR Explorer

Submissions API

Background

Via EDGAR, the SEC provides submissions history and the XBRL data from financial statements (forms 10-Q, 10-K, 8-K, 20-F, 40-F, 6-K, and their variants) for a company (filer) via a Representational State Transfer (REST or RESTful) application programming interface (API).[1] Using these services, users can download data about a company's filings by using the various APIs which include:

- Submissions API: the recorded history of all filings for a company by form name, date, and time of filing, etc. This API can be used to gather information on all the filings made by a single company. You will access the submissions API in this case "EDGAR Explorer Submissions API."
- XBRL data:
  - XBRL Company Concepts API: the XBRL disclosures from a single company and concept (a taxonomy and a tag). For example, you could use this API to gather information on Accounts Receivable for a single firm over several years.
  - XBRL Company Facts API: all the company concepts XBRL data, which includes all the XBRL data contained with the financial statements. For example, you could use this API to collect full financial statement information. You can access and explore the Company Facts API in the related EYARC Analytic Mindset Case "EDGAR Explorer Company Facts API."
  - XBRL Frames API: aggregation of one company fact most recently filed based on a requested calendar period of time (e.g., annual, quarterly). For example, you could use this API to gather information for accounts receivable for the most recent period.

The following information outlines some common features of these APIs:

- All the SEC's APIs can be accessed via a browser or automated using various programming languages or automation software.
- The SEC's APIs do not require any form of authentication or API keys to access the data. Use of the API is, however, subject to the SEC's web site privacy and security policy. In this case, we are required to follow the SEC's programmatic downloads requirement and declare a User-Agent in our requests.
  - A User-Agent header can be declared using the name of your University and University email address.
  - Users can also only make 10 requests per second.
- The SEC's APIs return data in JavaScript Object Notation (JSON) format. According to Wikipedia, "JSON is an open standard file format and data interchange format that uses human-readable text to

[1] The SEC's API documentation is available at: https://www.sec.gov/edgar/sec-api-documentation

store and transmit data objects consisting of attribute–value pairs and arrays (or other serializable values)."[2]

  - The SEC's JSON files use a combination of objects (key-value pairs) and arrays (lists of data). In this format, the SEC JSON files will include metadata, such as the company's central index key (CIK) and company name.
  - Note that public companies will generally report line items both from the standard XBRL taxonomy as well as extended taxonomies that can be specific to a company and/or industry.

- The SEC also provides access to XBRL tagged financial data in several other formats. For example, the SEC provides the Financial Statements and Notes Datasets which contain XBRL data for a specific period (one month since October 2020 and quarterly before then back to Q1 2009) in tab separated value (TSV) format.[3] These files are useful for broad comparisons of entities across the entire market.

Submissions API

This case uses the SEC's EDGAR Submissions API which is accessed via https://data.sec.gov/submissions/CIK##########.json by replacing the 10 hashes with the entity's unique 10-digit central index key (CIK). For example, using Tesla's CIK "0001318605" the submissions file for Tesla can be found at:

- https://data.sec.gov/submissions/CIK0001318605.json

The SEC's EDGAR website maintains a database of company names and tickers to help users look up entities' CIKs to access their filings. It can be found at:

- https://www.sec.gov/edgar/searchedgar/cik.htm

The submissions file contains both metadata about the entity (including CIK, company name, geographical and industry data) and a large amount of data describing the various disclosures made by the entity. This case explores the types and frequencies of disclosures made by a specific public company. Some of the information provided includes:

- The accension number (`accessionNumber`) which can be used to locate the original full-text filing on EDGAR.
- Information about the day and time of the filing, including the date of the filing (`filingDate`), the filing report date (`reportDate`) and the date and time the filing was accepted by the SEC (`acceptanceDateTime`).
- Information about the type of form filed with the SEC, including the form type (`form`), and when applicable, the items (`items`) included in the form and/or which Act (`act`) the filing is required by (e.g., the 1934 Securities and Exchange Act). Additional descriptive information about the filings including the size of the filing and whether it is an XBRL filing (`isXBRL` and `isInlineXBRL`) are included along with two variables describing the primary document related to the filing (`primaryDocument` and `primaryDocDescription`).

[2] https://en.wikipedia.org/wiki/JSON
[3] https://www.sec.gov/dera/data/financial-statement-and-notes-data-set.html

- Important form types include the 10-K and 10-Q filings which contain financial statements, however, there are many file types that are more commonly seen for most companies including 8-K filings (the current report) and Form 4 filings (insider transactions).
- The 8-K filing reports on current material events, which is further classified into various types of events using an item number. For example, Item 4.01 relates to changes in the entity's auditor and Item 9.01 relates to the provision of financial statements and exhibits which often accompany an earnings announcement.

- Finally, at the end of the submission form, if the entity has filed more than 1,000 filings, there is information on earlier filings. In our Tesla example, the file references an earlier file as:
  - `"files":[{"name":"CIK0001318605-submissions-001.json","filingCount":265,"filingFrom":"2005-02-17","filingTo":"2012-12-17"}]`
  - Note that as Tesla makes future filings, the submission file on the API will push older filing information into the prior submission's JSON file. This means that the filing to date and count will change in the current file.

Required

- Use Tesla's CIK "0001318605" or pick a company and find its CIK using the EDGAR Company Filings CIK Lookup Tool at https://www.sec.gov/edgar/searchedgar/companysearch.
- Using Alteryx:
  - Use the download tool to download the submissions filing for your company. Remember to declare your University Name and University email address in the User-Agent header.
  - Deserialize the JSON formatted submissions filing for your company.
  - Use the output to
    - Identify the number of unique forms filed by the entity.
    - Calculate the number of times each form was filed based in this dataset and identify the most popular filing for your company.
- The following questions require researching forms filed with the SEC:
  - Based on your analysis above, pick two forms you have not heard of before, and research the type of forms that have been filed by your example company. The SEC website https://www.sec.gov/forms provides a useful starting point for your research (Note that a filing with /A appended to the end is an amended filing. Amended filings are corrected filings that have been previously filed with the SEC). Write a brief description of the forms you are researching and when describing the forms try to identify:
    - (i) what information is being disclosed by the company, and
    - (ii) why this information is expected to be important to users of financial information.
- Return to your Alteryx workflow (or python script) and use the appropriate tools to identity the most recent 10-K and use the `accessionNumber` with the SEC EDGAR Archive data URL: "https://www.sec.gov/Archives/edgar/data" and the company's CIK to download the most recent 10-K and save it as a csv file. Note that the website address that resolves the filing will be a text document, with the extension ".txt"
- Save your Alteryx workflow and submit for grading.