

Explore

This document provides information about the Microsoft Database Application and related Excel extraction tools which you can experiment with.

This information can be found here:

<http://xbrlsite.azurewebsites.net/2020/master/explore/index.html>

This database contains 109,778 US GAAP-based financial reports submitted to the SEC over a period of 10 years by 3,600 different public companies that used 17 different reporting styles. This subset of reports is used because (a) it is proven to be the case that fundamental information can be extracted from the report (i.e. "Assets" and "Liabilities and Equity"), (b) the rule "Assets = Liabilities and Equity" is proven to be adhered to, and (c) reports exist for at least 5 financial reporting periods for each economic entity. As such, the sum of "Assets" agrees with the sum of "Liabilities and Equity" for all 109,778 reports and 3,600 reporting entities. There is a difference of \$20 which relates to a handful of accumulated rounding errors. (Run the query qryTest_Balances)

SumOfAssets	SumOfLiabilitiesAndEquity	Difference	Count
1.17157122301665E+15	1.17157122301663E+15	20	109778

When using this database that points to reports you can either use each of the 109,778 reports online from the SEC web site or you can download all of the reports and cache them locally. Downloading all reports takes about 24 hours. Querying the reports locally cuts down on query time, it takes about 3 hours to query all the reports. Loading all of the reports into some sort of database cuts query time down even more; I have seen the query of assets and liabilities and equity taking as little as 22 seconds (MongoDB database).

The following is a breakdown of the reports that were submitted to the SEC but were excluded from this database:

1. All reports submitted using the 2009 US GAAP XBRL Taxonomy were excluded because there were so many errors. Basically, every public company was given this Mulligan.
2. If a reporting entity had an error caused by the document period end date which causes issues with finding any information in the report; then 100% of those reports of that economic entity were excluded and that economic entity did not get put into the database application.
3. If a reporting entity had an error related to finding the default economic entity (i.e. typically there is an XBRL dimensions use related problem); then 100% of the reports for that economic entity were excluded and that economic entity did not get put into the database application.
4. Only the top 17 reporting styles were used, if an economic entity did not use one of those most popular 17 reporting styles, then that economic entity and all of its reports were excluded from the database application.

So, what that leaves are 109,778 reports which are guaranteed to be readable by an XML parser of XBRL processor for 3,600 economic entities that report using 17 different reporting styles which have reported a combined total of 103,951,234 facts. The average number of facts per report is 946, the high

is 22,265 facts (SPIRIT REALTY CAPITAL, INC. 2016 10-K), the low was 44 facts (NATURAL GAS FUELING & CONVERSION INC. 2014 10-K).

Why is this set of information important? This information set is important because it allows the user of the database to explore the contents of the database and the reports the database points to where the fundamental logic of the reported information is intact. This allows for further exploration of detailed information that disaggregates “Assets”, “Liabilities and Equity”, “Net Income (Loss)”, and “Net Cash Flow”.

Additionally, there are five Excel spreadsheets that can be used to extract information from five of the 17 reporting styles used by reporting public companies. (I am not sure if I am going to build out the remaining 12, this depends on a few factors, I may reorganize the approach I am using.) The following is a summary of the 17 reporting style codes: (Excel extraction tools exist for the first five reporting styles indicated below in bold)

Reporting Style Code	Entities	Reports
COMID-BSC-CF1-ISM-IEMIB-OILY-SPEC6	1363	42,485
COMID-BSC-CF1-ISS-IEMIB-OILY-SPEC1	624	17,104
COMID-BSC-CF1-ISS-IEMIB-OILY-SPEC2	562	17,392
INTBX-BSU-CF1-ISS-IEMIX-OILN	371	11,828
INSBX-BSU-CF1-ISS-IEMIX-OILN	79	2,542
COMID-BSC-CF1-IS6-IEMIX-OILN	77	2,436
COMID-BSC-CF1-IS3-IEMIB-OILN	71	1,923
COMID-BSC-CF1-IS8-IEMIB-OILN	57	1,754
COMID-BSC-CF1-ISM-IEMIB-OILY-SPEC6-SCI2	55	1,683
COMID-BSC-CF1-ISM-IEMIT-OILY-SPEC6	52	1,688
COMID-BSC-CF1-ISS-IEMIB-OILY-SPEC2A	51	1,605
COMID-BSC-CF2-ISM-IEMIB-OILY-SPEC6	49	1,426
INTBX-BSU-CF1-ISS-IEMIX-OILN-SCI2	47	1,326
COMID-BSC-CF1-IS4-IEMIB-OILN	42	1,275
COMID-BSN-CF1-ISM-IEMIB-OILY-SPEC6	41	1,370
COMID-BSC-CF1-ISS-IEMIT-OILY-SPEC2	39	1,272
COMID-BSC-CF1-ISS-IEMIB-OILY-SPEC2-SCI2	20	668
Total	3,600	109,777

The Excel extraction tools cover about 80% of all the entities/reports in the database application.

The 17 reporting styles can be broken down into the following different styles of creating the primary financial statements:

- **Balance sheet:** 3 styles, classified, unclassified, and a classified alternative.
- **Cash flow statement:** 2 styles, exchange gains part of net cash flow and exchange gains as part of the cash and cash equivalents roll forward.
- **Comprehensive income statement:** 2 styles, where statement starts with Net Income (Loss) and where statement starts with Net income (Loss) Attributable to Parent.

- Income statement:** 9 styles, with gross profit, without gross profit, interest-based reporting, insurance-based reporting, with operating income (loss) explicitly reported, with operating income (loss) not explicitly reported, income (loss) from equity method investments reported as part of the tax provision, and so forth.

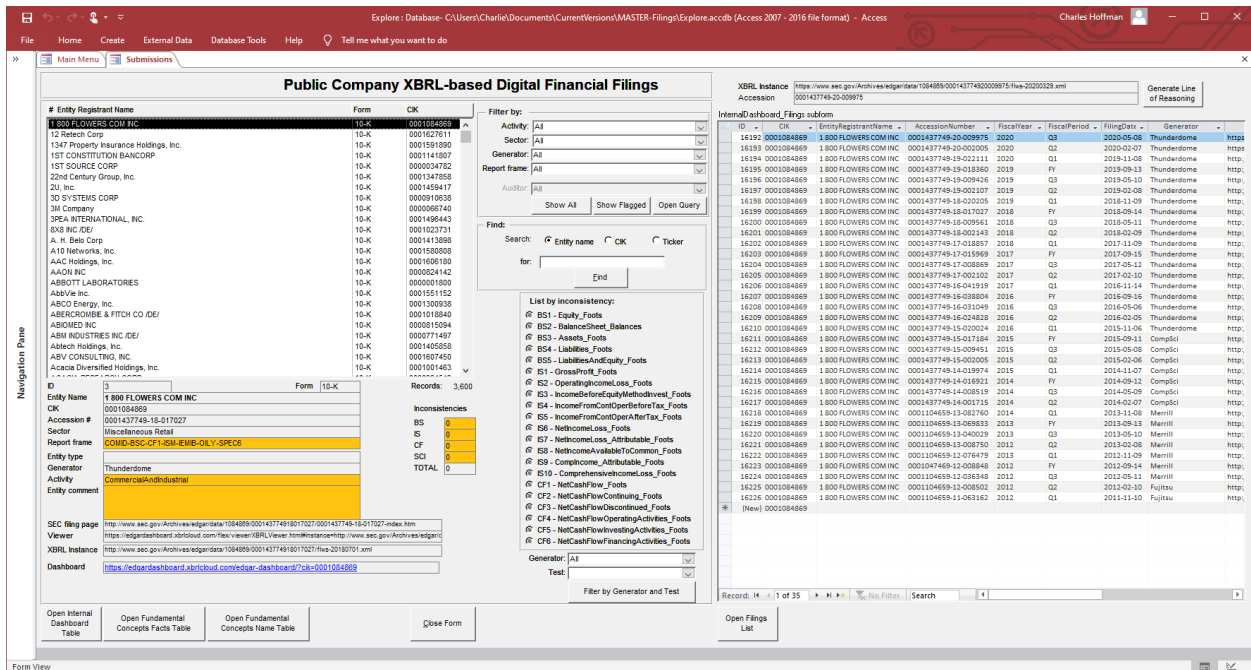
Application

The following provides information about the Microsoft Database application. The application has a table of entities (InternalDashboard), a table of reports for that entity (InternalDashboard_Filings), and a table with the results of extracting Assets, LiabilitiesAndEquity, a few other facts, and checking the difference between Assets and LiabilitiesAndEquity (Proof).

When the application starts a main form is loaded (frmMain). From the main form you can browse the entities (frmSubmissions), filter entities in a number of ways, and get a list of reports for the selected entity.

A user of this application has all the functionality of Microsoft Access to filter, sort, and otherwise slice and dice the lists of entities and reports.

Primary interface form:



From the primary interface, a user can select an entity, press the button “Open Filings List”, and get a query which contains a list of reports for the selected entity:

ID	CIK	EntityRegistrantName	AccessionNumber	FiscalYear	FiscalPeriod	FilingDate	Generator	
16192	0001084869	1 800 FLOWERS COM INC	0001437749-20-009975	2020	Q3	2020-05-08	Thunderdome	https
16193	0001084869	1 800 FLOWERS COM INC	0001437749-20-002005	2020	Q2	2020-02-07	Thunderdome	https
16194	0001084869	1 800 FLOWERS COM INC	0001437749-19-022111	2020	Q1	2019-11-08	Thunderdome	http;
16195	0001084869	1 800 FLOWERS COM INC	0001437749-19-018360	2019	FY	2019-09-13	Thunderdome	http;
16196	0001084869	1 800 FLOWERS COM INC	0001437749-19-009426	2019	Q3	2019-05-10	Thunderdome	http;
16197	0001084869	1 800 FLOWERS COM INC	0001437749-19-002107	2019	Q2	2019-02-08	Thunderdome	http;
16198	0001084869	1 800 FLOWERS COM INC	0001437749-18-020205	2019	Q1	2018-11-09	Thunderdome	http;
16199	0001084869	1 800 FLOWERS COM INC	0001437749-18-017027	2018	FY	2018-09-14	Thunderdome	http;
16200	0001084869	1 800 FLOWERS COM INC	0001437749-18-009561	2018	Q3	2018-05-11	Thunderdome	http;
16201	0001084869	1 800 FLOWERS COM INC	0001437749-18-002143	2018	Q2	2018-02-09	Thunderdome	http;
16202	0001084869	1 800 FLOWERS COM INC	0001437749-17-018857	2018	Q1	2017-11-09	Thunderdome	http;
16203	0001084869	1 800 FLOWERS COM INC	0001437749-17-019569	2018	FY	2017-09-15	Thunderdome	http;

The list of reports can be copied and then pasted into one of five Excel spreadsheet applications into the “List” sheet.

A	B	C	D	E	F	G	H	I	J
CIK	EntityRegistrantName	LinkToXBRLInstance	Generator	AccessionNumber	ReportingStyle	Form	FilingDate	FiscalYear	FiscalPeriod
0001506251	Citius Pharmaceuticals, Inc.	https://www.sec.gov/Archives/edgar/data/1506251/000121390020012286/ctxr-20200331.xml	GoXBRL	0001213900-20-012286	COMID-BSC-CF1-ISS-EMIB-OILY-SPECL	10-Q	2020-05-14	2020	Q2
0001506251	Citius Pharmaceuticals, Inc.	https://www.sec.gov/Archives/edgar/data/1506251/00012139002003577/ctxr-20191231.xml	GoXBRL	0001213900-20-003577	COMID-BSC-CF1-ISS-EMIB-OILY-SPECL	10-Q	2020-02-13	2020	Q1
0001506251	Citius Pharmaceuticals, Inc.	https://www.sec.gov/Archives/edgar/data/1506251/000121390019026229/ctxr-20190930.xml	GoXBRL	0001213900-19-026229	COMID-BSC-CF1-ISS-EMIB-OILY-SPECL	10-K	2019-12-16	2019	FY
0001506251	Citius Pharmaceuticals, Inc.	https://www.sec.gov/Archives/edgar/data/1506251/000121390019015793/ctxr-20190630.xml	GoXBRL	0001213900-19-015793	COMID-BSC-CF1-ISS-EMIB-OILY-SPECL	10-Q	2019-06-14	2019	Q3
0001506251	Citius Pharmaceuticals, Inc.	https://www.sec.gov/Archives/edgar/data/1506251/000121390019008833/ctxr-20190331.xml	GoXBRL	0001213900-19-008833	COMID-BSC-CF1-ISS-EMIB-OILY-SPECL	10-Q	2019-05-15	2019	Q2
0001506251	Citius Pharmaceuticals, Inc.	https://www.sec.gov/Archives/edgar/data/1506251/000121390019002563/ctxr-20181231.xml	GoXBRL	0001213900-19-002563	COMID-BSC-CF1-ISS-EMIB-OILY-SPECL	10-Q	2019-02-14	2019	Q1
0001506251	Citius Pharmaceuticals, Inc.	https://www.sec.gov/Archives/edgar/data/1506251/000121390018017214/ctxr-20180930.xml	GoXBRL	0001213900-18-017214	COMID-BSC-CF1-ISS-EMIB-OILY-SPECL	10-K	2018-12-11	2018	FY
0001506251	Citius Pharmaceuticals, Inc.	https://www.sec.gov/Archives/edgar/data/1506251/000121390018011072/ctxr-20180331.xml	GoXBRL	0001213900-18-011072	COMID-BSC-CF1-ISS-EMIB-OILY-SPECL	10-K	2018-08-11	2018	FY

The user can go to the “Compare” sheet, press the button “Compare All in List Spreadsheet” which will extract approximately 60 different facts from the report and test about 22 relations between those facts:

The screenshot shows an Excel spreadsheet titled "Compare" with a grid of financial data. The columns are labeled with various financial metrics such as Net Cash Flows, Earnings, and Retention Rates. The rows represent different reporting periods. The data is organized into sections, with a "Key Financial Ratios" section highlighted in yellow. The spreadsheet is used for comparing and testing relations between different financial facts.

Be sure to match the spreadsheet you use and the reporting style code otherwise the extraction algorithms will not work correctly.