The Annual Report Algorithm:

Retrieval of Financial Statements and Extraction of Textual Information*

Jörg Hering[1]

Friedrich-Alexander-Universität Erlangen-Nürnberg

First Version: October 4, 2016

Current Version: November 28, 2016

[1] Jörg Hering, Friedrich-Alexander-Universität Erlangen-Nürnberg, School of Business and Economics, Department of Accounting and Auditing, Lange Gasse 20, 90403 Nuremberg, Germany, phone: +49 911 5302 341, e-mail: joerg.hering@fau.de.

* I appreciate the valuable comments by Klaus Henselmann and Daniel Büchs.

The Annual Report Algorithm:

Retrieval of Financial Statements and Extraction of Textual Information

Abstract

U.S. corporations are obligated to file financial statements with the U.S. Securities and Exchange Commission (SEC). The SEC´s Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system containing millions of financial statements is one of the most important sources of corporate information available. The paper illustrates which financial statements are publicly available by analyzing the entire SEC EDGAR database since its implementation in 1993. It shows how to retrieve financial statements in a fast and efficient way from EDGAR. The key contribution however is a platform-independent algorithm for business and research purposes designed to extract textual information embedded in financial statements. The dynamic extraction algorithm capable of identifying structural changes within financial statements is applied to more than 180,000 annual reports on Form 10-K filed with the SEC for descriptive statistics and validation purposes.

Keywords: Textual analysis, Textual sentiment, 10-K parsing rules, Information extraction, EDGAR search engine

1. Introduction

Information Extraction (IE) can be defined as the process of *"finding and extracting useful information in unstructured text"* (Grant and Conlon 2006, 119). In contrast to Information Retrieval (IR), a technology that selects a relevant subset of documents from a larger set, IE extracts information from the actual text of documents (Wilks 1997, 1). Important sources for IE are unstructured natural language documents or structured databases (Mooney and Bunescu 2005, 3; Gaizauskas et al. 1997, 28). Since U.S. corporations are obligated by law to file financial statements on a regular basis with the U.S. Securities and Exchange Commission (SEC), the SEC´s Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system containing millions of financial statements is one of the most important sources of corporate information available (Garcia and Norli 2012, 2; Grant and Conlon 2006, 120).

Unfortunately, most of the available textual data in the SEC EDGAR database is weakly structured in technical terms (Stümpert 2008, 357-364; Bovee et al. 2005, 25; O´Riain 2012, 35) especially prior to 2002 when the use of markup languages was less common (Loughran and McDonald 2014, 1650). A limited number of tagged items, formatting errors and other inconsistencies lead to difficulties in accurately identifying and parsing common textual subjects across multiple filings (Gerdes 2003, 16-17; Kambil and Ginsburg 1998, 92-93; Bovee et al. 2005, 20). These issues directly affect the ability to automate the extraction of textual information from SEC submissions (Gerdes 2003, 16; Davis and Tama-Sweet 2012, 811; Loughran and McDonald 2016, 1191-1192). Business data providers are offering expensive commercial products (e.g. AcademicEDGAR+, Edgar Pro, Intelligize). As research in the context of textual analysis is growing (e.g. Tetlock 2007; Loughran and McDonald 2011a; Jegadeesh and Wu 2013) the question occurs which particular financial statements and disclosures are publicly available for free, how to retrieve these corporate documents and how to decode the embedded textual information in order to be incorporated into investment decisions, trading strategies and research studies in financial economics (Garcia and Norli 2012, 1).

Today only a very limited amount of specific literature for extracting textual information from financial statements filed with the SEC and its EDGAR system is available (except Gerdes 2003; Stümpert et al. 2004; Grant and Conlon 2006; Engelberg and Sankaraguruswamy 2007; Cong, Kogan and Vasarhelyi 2007; Thai et al. 2008; Chakraborty and Vasarhelyi 2010; Hernandez et al. 2010; Garcia and Norli 2012; Srivastava 2016). This paper is based on neither of these because first, non-specialist technology is used to retrieve financial statements in an efficient way and secondly, the algorithm designed to extract textual information is platform-independent. The suggested method can compensate for expensive commercial products and help to replicate empirical research results. The paper shall serve as a technical guide on how to retrieve financial statements filed with the SEC and

how to decode the embedded textual information provided by the EDGAR system for business and research purposes.

The remainder of the paper proceeds as follows. Section 2 presents the amount and variety of corporate documents distributed by the SEC´s electronic disclosure system. Section 3 demonstrates how to retrieve these documents from the EDGAR database. Section 4 describes the fundamentals of HyperText Markup Language and examines the electronic data provided by the SEC. Section 5 describes the fundamentals of regular expressions and specifies an algorithm to extract textual information embedded in financial statements. Section 6 validates the capabilities of the extraction algorithm. Section 7 presents descriptive statistics of annual reports filed with the EDGAR database. The last section concludes.

2. SEC´s EDGAR database

Publicly owned companies, their officers and directors as well as major investors are obligated by law (Securities Exchange Act 1934, Section 2) to file various disclosures (forms) with the SEC (Gerdes 2003, 7). The main purpose of making certain types of corporate information publicly available is to improve the efficiency of security markets and to protect capital market participants (Garcia and Norli 2012, 2). *"The laws and rules that govern the securities industry in the United States derive from a simple and straightforward concept: all investors, whether large institutions or private individuals, should have access to certain basic facts about an investment prior to buying it, and so long as they hold it. To achieve this, the SEC requires public companies to disclose meaningful financial and other information to the public. This provides a common pool of knowledge for all investors to use to judge for themselves whether to buy, sell, or hold a particular security"* (SEC 2013). In order to protect investors, to maintain efficient capital markets and to improve access to publicly available corporate disclosures, the SEC developed the EDGAR database (Gerdes 2003, 9) and describes it as a system which *"performs automated collection, validation, indexing, acceptance, and forwarding of submissions by companies and others who are required by law to file forms with the U.S. Securities and Exchange Commission"* (SEC 2010).

Originally the EDGAR system was developed by the SEC as a pilot system for electronic disclosure in 1983. In order to test and evaluate EDGAR´s performance the SEC requested electronic filings in 1994 after completing the phase-in of a mandated test group in December 1993 (the phase-in began on April 26, 1993) (SEC 2006; Kambil and Ginsburg 1998, 91; Pagell 1995, 56). As of May 6, 1996 the SEC obligated all public domestic U.S. companies (issuers) to file submissions electronically through the EDGAR system (SEC Release 34-36997 1996; Kambil and Ginsburg 1998, 91; Pagell 1995, 56; Grant and Conlon 2006, 121) except for certain filings made in paper because of a hardship exemption under Regulation S-T (SEC Regulation S-T 2016, Section 232.201; SEC,

2010). Filing for foreign private issuers (companies organized outside of the U.S.) and foreign governments via EDGAR (SEC 2006) became mandatory on May 14, 2002 (SEC Release 33-8099 2002). The Securities Exchange Act of 1934 (Securities Exchange Act 1934, Section 13(a), (b), Section 15(d)) empowers the SEC to require (periodic) reporting of information from publicly held companies (SEC 2013). In general, all public domestic companies with assets exceeding $10 million and at least 500 shareholders become subject to Exchange Act reporting requirements (Securities Exchange Act 1934, Section 12(g)) alongside certain individuals (Gerdes 2003, 9). Among other disclosures, corporations with publicly traded securities are required (Securities Exchange Act 1934, Section 13(a), (b), Section 15(d)) to file annual and quarterly reports (Form 10-K, Form 10-Q) as well as current reports (Form 8-K) on an ongoing basis with the SEC and its EDGAR system (SEC 2013). Since by law these public corporate disclosures have to be accurate (Securities Exchange Act 1934, Section 13(i)) and represent a company´s operations, they themselves represent a treasure trove of valuable information for investors and researchers (Gerdes 2003, 9; Engelberg and Sankaraguruswamy 2007, 3).

2.1 Underlying data in SEC´s EDGAR database

In order to understand the amount and variety of corporate information (e.g. financial statements) distributed by the SEC, I retrieve and analyze all form index files since the implementation of the EDGAR system in 1993. The SEC EDGAR form index files list all publicly available disclosures made through the system in a certain quarter and sort the submissions by their particular filing form type. Table 1 reports the total number of submissions that have been made with the EDGAR system for each quarter and year since the introduction of the EDGAR database.

Table 1. Statistics on EDGAR submissions.

Year	**Filings** (Number)				**Filings** (Number)	**Filings** (%)
	Quarter 1	Quarter 2	Quarter 3	Quarter 4		
2016	307,416	239,528	---	---	546,944	3.42
2015	318,519	259,852	206,628	209,216	994,215	6.21
2014	311,679	252,333	212,352	220,328	996,692	6.23
2013	303,568	257,597	213,031	216,266	990,462	6.19
2012	309,453	246,776	203,723	214,985	974,937	6.09
2011	307,644	262,218	207,142	202,628	979,632	6.12
2010	300,538	255,180	203,920	220,070	979,708	6.12
2009	300,080	229,347	200,688	208,396	938,511	5.87
2008	328,709	267,722	220,732	219,669	1,036,832	6.48
2007	339,872	289,082	252,071	256,460	1,137,485	7.11
2006	335,577	278,960	232,131	249,956	1,096,624	6.85
2005	317,761	271,632	242,173	240,725	1,072,291	6.70
2004	312,029	253,021	217,726	241,435	1,024,211	6.40
2003	183,595	167,119	212,258	227,800	790,772	4.94
2002	125,189	108,013	97,533	118,149	448,884	2.81
2001	111,740	90,283	74,313	75,107	351,443	2.20
2000	116,209	81,129	72,571	72,053	341,962	2.14
1999	105,531	78,272	68,631	68,828	321,262	2.01
1998	106,666	73,830	67,234	65,570	313,300	1.96
1997	91,096	65,470	60,142	63,422	280,130	1.75
1996	49,925	47,659	50,641	54,389	202,614	1.27
1995	31,875	26,104	26,699	28,973	113,651	0.71
1994	20,879	16,500	13,066	15,016	65,461	0.41
1993	4	4	7	20	35	0.00
Filings (Number)	**5,035,554**	**4,117,631**	**3,355,412**	**3,489,461**	**15,998,058**	**100.00**
Filings (%)	**31.48**	**25.74**	**20.97**	**21.81**	**100.00**	

Notes: The table presents the total number of filings made on EDGAR for each year between 1993 and 2016. Each individual filing in a particular quarter is listed in an associated EDGAR form index file on the SEC server.

A tremendous amount of publicly available disclosures was filed with the SEC between 1993 and 2016. In total 15,998,058 filings were submitted to the EDGAR system in order to be publicly distributed. On average 31.48 percent (5,035,554) of these filings became available in the first, 25.74 percent (4,117,631) in the second, 20.97 percent (3,355,412) in the third and 21.81 percent (3,489,461) in the last quarter of each year since 1993. Most noticeable is the overall increase in total submissions through the EDGAR system reaching its peak in 2007 with more than 1.1 million disclosures for that particular year. By analyzing the index files more precisely, investors and researchers can gain an insight into the specific type of information the SEC is making publicly available through its EDGAR system (Garcia and Norli 2012, 2). Table 2 describes the most common filing (form) types filed with the EDGAR system.

Table 2. Statistics on EDGAR form types.

Rank	Form/Description	Submission Type	**Filings** (Number)	**Filings** (%)
1	Changes in ownership	4	5,850,937	36.57
2	Current report filing	8-K	1,376,248	8.60
3	5% passive ownership triggers amendments	SC 13G/A	587,711	3.67
4	Initial ownership report	3	538,228	3.36
5	Quarterly report	10-Q	522,906	3.27
6	Definitive materials	497	365,987	2.29
7	5% passive ownership triggers	SC 13G	344,030	2.15
8	Current report of foreign issuer	6-K	326,751	2.04
9	Change on a prospectus	424B3	254,046	1.59
10	5% active ownership triggers amendments	SC 13D/A	201,938	1.26
11	Changes in ownership amendments	4/A	197,612	1.24
12	Quarterly holdings, institutional managers	13F-HR	193,463	1.21
13	Annual report on ownership changes	5	186,884	1.17
14	Annual report	10-K	167,599	1.05
15	SEC-originated letters to filers	UPLOAD	159,065	0.99
16	Filer response letters	CORRESP	153,987	0.96
17	Proxy statements	DEF 14A	152,216	0.95
18	Registration management investment companies	485BPOS	151,903	0.95
19	Registration of securities, investment companies	24F-2NT	149,385	0.93
20	Offering of securities without registration	D	147,355	0.92
...	...	...	...	...
Total			**15,998,058**	**100.00**

Notes: The table presents the most frequent form types filed with the EDGAR system between 1993 and 2016. The first column ranks each filing type in descending order of total submissions. The second column gives a short description of each filing form type (Garcia and Norli 2012, 2). The third column lists the form codes used on EDGAR to identify a particular filing type made with the database. The next column contains the number of total submissions of a particular filing form type. The last column shows the amount of total submissions for each filing type in relation to all submissions made with the SEC EDGAR database.

The submission type most often filed with the EDGAR system since its implementation is Form 4. Between 1993 and 2016 5,850,937 filings report purchases or sales of securities by persons who are the beneficial owner of more than ten percent of any class of any equity security, or who are directors or officers of the issuer of the security (Garcia and Norli 2012, 2). The second most frequent submission type filed with the SEC is Form 8-K. 1,376,248 filings of this submission type are listed in the EDGAR index files. The current report filing is required by companies in order to inform shareholders about certain corporate events. These events of material importance for a company include information on significant agreements, impairments, changes in management etc. (Garcia and Norli 2012, 2). Important submission types for investors and researchers such as the annual report on Form 10-K have been submitted 167,599 times. Quarterly reports on Form 10-Q have been filed 522,906 times in total between 1993 and 2016. Another important submission type is Schedule 13G (SC 13G). Investors who are not seeking control over a firm (passive investors) must file this submission type as required by the SEC when crossing the five percent ownership threshold of a company (Garcia and Norli 2012, 2). In total 344,030 filings of this particular submission type alone are reported on EDGAR. Appendix B shows a detailed overview of all publicly available submissions made with the SEC between 1993 and 2016.

The SEC assigns to each filer a Central Index Key (CIK) which is a unique identifier used on the EDGAR database in order to label and identify each individual filer in the system (Gerdes 2003, 16). Since 1993 in total 580,225

unique CIK numbers were assigned and stored in the SEC´s electronic disclosure system. The majority of these CIKs were not assigned to publicly traded companies but to private firms, hedge funds and mutual funds as well as to private individuals who receive a CIK when filing with the SEC (Garcia and Norli 2012, 2). Table 3 reports the number of unique CIKs (unique filers) filing a certain submission type with the SEC and its EDGAR system.

Table 3. Statistics on EDGAR filers.

Rank	Form/ Description	Submission Type	Unique CIKs	Mean	Med.	Max.
1	Changes in ownership	4	206,652	28.3	7	12,170
2	Initial ownership report	3	187,366	2.9	1	550
3	Offering of securities without registration	D	104,853	1.4	1	375
4	Regulation D exemption filing (paper submission)	REGDEX	87,285	1.5	1	150
5	Changes in ownership amendments	4/A	62,099	3.2	1	338
6	Annual report on ownership changes	5	47,466	3.9	1	473
7	Change on a prospectus	424B3	45,204	5.6	2	9,911
8	5% active ownership triggers	SC 13D	43,381	2.3	1	730
9	5% passive ownership triggers	SC 13G	41,629	8.3	2	7,726
10	Notification of effectiveness for Securities Act registration statement	EFFECT	40,485	2.4	1	86
11	Registration of securities issued in business combination transactions	S-4	40,139	2.0	1	70
12	Current report filing	8-K	38,713	35.6	10	1,484
13	Offering of securities without registration amendments	D/A	35,673	2.8	2	1,601
14	Registration of securities issued in business combination transactions amendments	S-4/A	35,158	2.8	2	63
15	Annual report	10-K	33,968	4.9	3	107
16	5% passive ownership triggers amendments	SC 13G/A	33,339	17.6	4	25,447
17	SEC-originated letters to filers	UPLOAD	31,720	5.0	3	91
18	Filer response letters	CORRESP	30,031	5.1	3	157
19	5% active ownership triggers amendments	SC 13D/A	29,742	6.8	3	5,528
20	Quarterly report	10-Q	26,271	19.9	14	189

Notes: The table presents the most frequent submission types made on EDGAR in descending order of unique SEC registrants filing a particular submission type. The time period is 1993-2016. The fourth column contains the total number of unique filers submitting a particular form type. Columns 5-7 present the means, medians and maxima of particular filing form types submitted by unique SEC filers.

Submission type Form 4 (Form 3) was submitted by 206,652 (187,366) different filers between 1993 and 2016. Annual reports on Form 10-K were submitted to the SEC by 33,968 filers. Quarterly reports on Form 10-Q can be associated with 26,271 unique filers whereas the number of CIKs assigned to current reports on Form 8-K is 38,713. On average each registrant filed 4.9 annual reports on Form 10-K and 19.9 quarterly reports on Form 10-Q with the EDGAR system in addition to 35.6 current reports on Form 8-K since 1993. AFS SenSub Corp. (CIK 1347185), an issuer of asset-backed securities, filed 107 annual reports on Form 10-K (56 on 10-K/A). PowerShares DB Multi-Sector Commodity Trust (CIK 1367306), an investment company offering several investment funds, filed 189 quarterly reports on Form 10-Q (7 on 10-Q/A). Chase Bank USA, National Association (CIK 869090) filed 1,484 Form 8-K statements (12 on 8-K/A). 730 Schedule 13D Forms were filed by Gamco Investors, INC. (CIK 807249), an investment advisory and brokerage service firm, (5,528 on SC 13D/A) whereas

FMR LLC (CIK 315066), the financial services conglomerate known as Fidelity Investments, filed 7,726 Schedule 13G Forms (25,447 on SC 13G/A).

3. SEC EDGAR Data Gathering

Researchers in the field of finance and accounting often rely on programming languages (Perl, Python, R, SAS, and SPSS) to retrieve financial statements filed with the SEC. The use of a programming language as a tool is problematic for several reasons. First, many people analyzing financial reports are not familiar with these programming languages. For them it is time-consuming to apply a specific and complex coding language to obtain the corporate filings from EDGAR. Secondly, due to downloading only one filing at a time the procedure is very slow especially when obtaining massive data from the database. Thirdly, since incremental changes have to be made to the algorithm to retrieve another filing form type or filings from another company this particular method is very error-prone.

In contrast, widely used internet browsers (e.g. Mozilla-Firefox, Google-Chrome) can be easily equipped with powerful applications (e.g. DownThemAll, GetThemAll) which offer advanced download capabilities. These fully integrated browser extensions are able to identify links contained in a webpage or file and download the desired document parts simultaneously. To feed these applications only a standard MS Excel spreadsheet is necessary.

Every filing made through the EDGAR system in a particular quarter between 1993 and 2016 is stored in an associated index file (file extension *.idx) (Garcia and Norli 2012, 3). The EDGAR index files therefore represent a helpful resource in retrieving massive data from the database. They list important information for each filing such as the name of the filer, the particular central index key, the date and the type of the submission as well as the particular name of the document on the SEC server. In general, four different types of index files are available sorting the filings made on EDGAR by company name, form type, central index key or by submissions containing financial statements formatted in eXtensible Business Reporting Language (XBRL)[1] (SEC 2015). Appendix A lists all index files sorting the entire EDGAR database by filing form type for every quarter between 1993 and 2016 (SEC Index Files 2016). When examining the form index files more precisely one can see that the index files do not only contain the name of any filing made on EDGAR but rather the (entire) server path. Table 4 illustrates an excerpt of information stated in the SEC EDGAR form index file from the first quarter of 2016.

[1] eXtensible Business Reporting Language (XBRL) is a meta language for creating markup languages. XBRL was created for the electronic exchange of business data (Ditter, Henselmann and Scherr 2011, 22). XBRL is a variant of the Extensible Markup Language (XML) and related to the HyperText Markup Language (HTML). XBRL provides semantic context for data reported in SEC EDGAR submissions (Form 10-K) (Bodnaruk, Loughran and McDonald 2015, 643).

Table 4. SEC EDGAR 2016 form index file.

Form	**Company Name**	**CIK**	**Date Filed**	**File Name** (including partial directory)
10-K	AMAZON COM INC	1018724	2016-01-29	edgar/data/1018724/0001018724-16-000172.txt
10-K	Alphabet Inc.	1652044	2016-02-11	edgar/data/1652044/0001652044-16-000012.txt
10-K	COCA COLA CO	21344	2016-02-25	edgar/data/21344/0000021344-16-000050.txt
...	...	...	...	...
10-Q	STARBUCKS CORP	829224	2016-01-26	edgar/data/829224/0000829224-16-000049.txt
10-Q	APPLE INC	320193	2016-01-27	edgar/data/320193/0001193125-16-439878.txt
10-Q	VISA INC	1403161	2016-01-28	edgar/data/1403161/0001403161-16-000015.txt
...	...	...	...	...
8-K	GOLDMAN SACHS INC	886982	2016-01-20	edgar/data/886982/0001193125-16-433035.txt
8-K	EBAY INC	1065088	2016-01-27	edgar/data/1065088/0001065088-16-000262.txt
8-K	MICROSOFT CORP	789019	2016-01-28	edgar/data/789019/0001193125-16-441813.txt
...	...	...	...	...

Notes: The table presents an excerpt of the EDGAR form index file from the first quarter of 2016. The first column shows the specific submission form type of each filing listed in the quarterly SEC EDGAR form index file. The second column shows the name of each EDGAR registrant submitting a filing. The Central Index Key (CIK) of each filer is shown in the next column. The last two columns contain the submission date and the document name of each filing on the EDGAR database.

By opening the index files for example with a simple MS Excel spreadsheet (file extension *.xlsx) a Uniform Resource Locator (URL) can be created for each financial statement which is listed in a particular index file since the name of the filing and its (partial) server path (directory) is stated. To do so the protocol (https://), the hostname (www.sec.gov/) and a link to the archives directory (Archives/) have to be added to the file name from the index file. Table 5 illustrates the URL components of Coca Cola´s 2015 annual report on Form 10-K filed with the SEC on February 25, 2016.

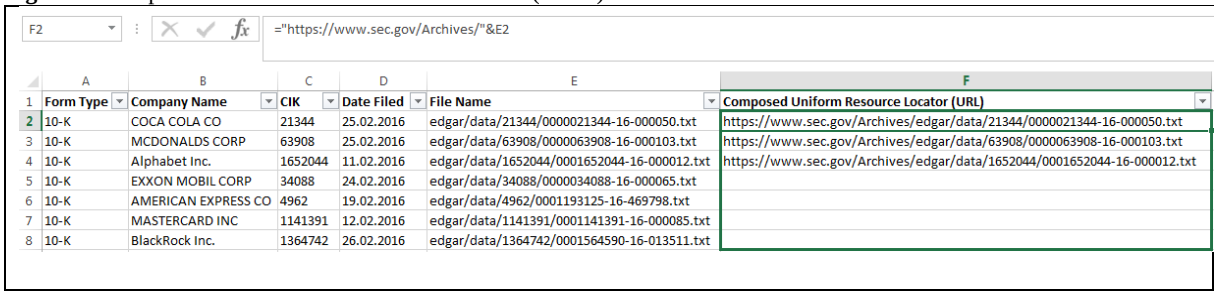
Table 5. Components of a Uniform Resource Locator (URL).

URL Component	**String**
Protocol	https://
Hostname	www.sec.gov
Directory	/Archives/edgar/data/21344/
File Name	0000021344-16-000050.txt
Composed URL	https://www.sec.gov/Archives/edgar/data/21344/0000021344-16-000050.txt

Notes: The table presents the different components of a Uniform Resource Locator. The first column shows the different component types within the URL. The second column shows the particular value of a certain component type in the URL.

Figure 1 illustrates the composition of URLs in MS Excel using a simple formula command.

Figure 1. Composition of Uniform Resource Locators (URLs) in MS Excel.



Notes: The table demonstrates the composition of URLs in MS Excel based on the SEC EDGAR form index files. The last column presents the URL for each disclosure which has been composed using the filename listed in the form index file and a simple MS Excel formula command (=”https://www.sec.gov/Archives/”&File Name).

These URLs which have been composed based on the EDGAR index files can be copied into a plain text file (file extension *.txt). By opening it with the browser extensive data (financial statements) can be retrieved from the SEC and its EDGAR system in a fast and efficient way using a browser extension (however, the composed URLs can also be implemented in any other data gathering method).

This method offers various significant advantages. First, for many people composing URLs with commonly used and easy accessible computer software like MS Excel is simpler and faster than relying on complex coding languages to identify and retrieve the documents in question. Secondly, since multiple documents can be retrieved at the same time using browser extensions, the described method is again a lot faster especially when obtaining massive data from EDGAR. Thirdly, by sorting or filtering the different index files in MS Excel the proposed method can easily be adjusted to retrieve another filing form type or data from another company. The result of this procedure is validated through obtaining exactly the same financial statements investors and researchers would retrieve using a complex, slow and error-prone alternative.

4. HyperText Markup Language in SEC filings

Because financial statements filed with the SEC are formatted in HyperText Markup Language (HTML) the fundamentals of HTML are illustrated first, followed by an examination of the data formatted in HTML provided by the SEC and its EDGAR system.

4.1 Fundamentals of HyperText Markup Language

HyperText Markup Language (HTML) is a universally understood digital language which is used to publish and distribute information globally. HTML is the publishing language of the World Wide Web (W3C Recommendation 1999, Section 2.2). HTML is used to create HyperText documents that are portable from one platform to another (Filer Manual 2016, Section 5-12) due to their generic semantics as a Standard Generalized Markup Language (SGML)[2] application (W3C Recommendation 1999, Section 3.1). HTML enables authors to publish documents online, assign a specific look or layout to document content (tagging) (Ditter, Henselmann and Scherr 2011, 20; Chakraborty and Vasarhelyi 2010, 4) or to retrieve information online via HyperText links (W3C Recommendation 1999, Section 2.2). The World Wide Web Consortium (W3C) is maintaining and specifying the

[2] Standard Generalized Markup Language (SGML) is a meta language or system for defining markup languages. A markup language defined in SGML is a *"SGML application"* (HTML) (W3C Recommendation 1999, Section 3.1).

vocabulary (applicable markups) and grammar (logical structure) of HTML documents (Ditter, Henselmann and Scherr 2011, 20).

A valid HTML document is composed of three different parts (W3C Recommendation 1999, Section 7.1). First, it declares which version of HTML is used in the document through the document type declaration (<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">). The document type declaration names the document type definition (DTD) specifying which elements and attributes can be implemented into a document formatted in HTML (W3C Recommendation 1999, Section 7.2). HTML 4.01 specifies three different DTDs: HTML 4.01 Strict DTD; HTML 4.01 Transitional DTD and HTML 4.01 Frameset DTD (W3C Recommendation 1999, Section 7.2). The W3C recommends to use HTML 4.01 Strict DTD which excludes presentation attributes since these elements are supposed to be replaced by style sheets (W3C Strict DTD, 1999). The second part of a HTML document is the document head (<HEAD>). This section contains information about the current document such as the title and relevant keywords for search engines. In general, the elements appearing in the head section are not presented by a document formatted in HTML (W3C Recommendation 1999, Section 7.4). The third and most important part of a HTML document is the body (<BODY>). This section contains the actual content of the document such as text paragraphs, images, graphics, tables, links, etc. (W3C Recommendation 1999, Section 7.5). The content in the document body can be structured in many different ways using various HTML elements (tags) to accomplish a certain look or layout to present the embedded information. Figure 2 illustrates a simplified excerpt of a document formatted in HTML (W3C Recommendation 1999, Section 7.1).

Figure 2. Document formatted in HTML 4.01.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">

<HTML>

  <HEAD>
    <TITLE>“Annual Report on Form 10-K”</TITLE>
  </HEAD>

  <BODY>“United States Securities and Exchange Commission …”</BODY>

</HTML>
```

Notes: The figure shows a simplified excerpt of a well-formed document formatted in HTML 4.01.

4.2 SEC EDGAR HTML Data

"Official" financial statements filed with the SEC have to be formatted either in American Standard Code for Information Interchange (ASCII) or in HyperText Markup Language (HTML 3.2/4.0).[3] Financial statements formatted in Portable Document Format (PDF) or XBRL are considered *"unofficial"* documents (submissions formatted in PDF and XBRL may qualify as official documents as well when specific criteria are met) (Filer Manual 2016, Section 5-1, Section 5-26). Due to a limited support of HTML in order to reduce the number of inconsistencies caused by HTML 4.0 implementation variances (SEC 2000), the EDGAR system only accepts a subset of HTML 3.2 semantics (tags) and several HTML 4.0 attributes (Filer Manual 2016, Section 5-17) therefore enforcing several restrictions (no active content, no external references etc.) of HTML formatting in financial statement submissions (Filer Manual 2016, Section 5-12).

The *"Complete Submission Text File"* (file extension *.txt) provided by the EDGAR system represents an aggregation of all information in a particular financial statement filed with the SEC. The text version of the filings on the SEC server contains the 10-K document formatted in HTML, XBRL, exhibits and ASCII-encoded graphics (*"binary-to-text"* encoding or *"uuencoding"* converts binary data files to plain ASCII-printable characters to facilitate transfer across various hardware platforms) (Bodnaruk, Loughran and McDonald 2015, 643; Loughran and McDonald 2011b, 1). Besides the *"Complete Submission Text File"* several submission parts (documents) are also provided in HTML (file extension *.htm) such as the core 10-K document and the exhibits which have been submitted (Bodnaruk, Loughran and McDonald 2015, 643). For example, Coca Cola´s 10-K filing on February 25, 2016 lists the core 10-K filing in HTML format, ten exhibits, eight graphic files (file extension *.jpg), six XBRL files and a single *"Complete Submission Text File"* containing all of these documents (SEC EDGAR Archives 2016). Figure 3 illustrates a simplified excerpt of Coca Cola´s 2015 annual report on Form 10-K formatted in HTML embedded in the *"Complete Submission Text File"*.

[3] American Standard Code for Information Interchange (ASCII) was created to standardize the storing of alphanumeric characters by defining a unique binary 7-bits number for each of the 128 (2^7) storable characters. ASCII included the Roman/Latin alphabet and essential characters for writing English. Using 8-bits (single 8-bit byte) the character encoding of HTML 2.0/4.01 is ISO-8859-1, an extension to ASCII which allows to present 256 (2^8) characters in HTML documents (W3 Schools, 2016; Palmer 2010, 11).

Figure 3. Coca Cola´s 2015 annual report on Form 10-K formatted in HTML 4.01.

```
<SEC-HEADER>…</SEC-HEADER>

<DOCUMENT>
<TYPE>10-K
<TEXT>
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">

<html>
  <head>
    <title>10-K</title>
  </head>

  <body style="font-family:Times New Roman;font-size:10pt;">…

  <div style...><font style...>UNITED STATES</font></div>…
  <div style...><font style...>SECURITIES AND EXCHANGE COMMISSION</font></div>…

  <div style...><font style...>FORM 10-K</font></div>...
  <div style...><font style...>For the fiscal year ended December 31, 2015</font></div>…

  <div style...><font style...>ITEM 1. BUSINESS</font></div>...
  <div style...><font style...>The Coca-Cola Company is the world's largest beverage company…

  <div style...><font style...>ITEM 1A. RISK FACTORS</font></div>...
  <div style...><font style...>In addition to the other information set forth in this report, you should…

  <div style...><font style...>ITEM 7. MANAGEMENT'S DISCUSSION AND ANALYSIS OF FINANCIAL
                               CONDITION AND RESULTS OF OPERATIONS</font></div>...
  <div style...><font style...>The following Management's Discussion and Analysis of Financial Condition…

  </body>

<html>

</TEXT>
</DOCUMENT>
```

Notes: The figure presents a simplified excerpt of Coca Cola´s 2015 annual report on Form 10-K formatted in HTML 4.01 contained in the *"Complete Submission Text File"* available on the SEC server.

5. Textual Information in Financial Statements

This section describes how regular expressions are used to extract textual information from financial statements filed with the SEC. First, I illustrate the fundamentals of regular expressions. Then I discuss the algorithm to extract textual information from financial statements using only regular expressions before presenting the actual text embedded in financial statements as a result of the designed algorithm. Due to their high relevance for investors and researchers an actual annual report on Form 10-K from the Coca Cola Company serves as basis for the illustration.

5.1 Fundamentals of Regular Expressions

Regular expressions or regular sets were first used as an algebra by mathematicians to describe models developed by physiologists of how the nervous system would work at the neuron level. The first published computational use of regular expressions was in 1968 by Ken Thompson (Friedl 2006, 85) who describes regular expressions as *"a*

method for locating specific character strings embedded in character text" (Thompson 1968, 419). They are implemented not only in modern programming languages, but also in application programs that can be used for text analysis without special programming skills (e.g. RapidMiner).

Regular expressions (*"RegEx"; "RegExp"; "RegExes"*) with a general pattern notation (pattern language) allow to process all kinds of text and data in a flexible and efficient way (Friedl 2006, 1; Loughran and McDonald 2016, 1219-1220). In particular RegExes can be used to modify textual elements or to identify and extract certain information from different documents (Goyvaerts and Levithan 2012, 1-2). The two types (full) regular expressions are composed of are special characters (metacharacters) and normal (literal) text characters acting as the grammar and the words of the regular expression language (Friedl 2006, 5; Goyvaerts and Levithan 2012, 28). For example, RegEx: *"[0-9]"* identifies all digits, RegEx: *"[a-zA-Z]"* isolates all upper and lower-case letters (character classes) and RegEx: *"."* matches all of these elements (metacharacter) embedded in an underlying text document (Friedl 2006, 9-12; Goyvaerts and Levithan 2012, 33-39). Another metacharacter and counting element (quantifier) within the regular expression language is a star or an asterisk (*) which quantifies the immediately preceding item within the defined expression (match any number of the preceding element including none) (Friedl 2006, 18; Goyvaerts and Levithan 2012, 74). Counting elements or quantifiers are used to specify the search pattern of regular expressions in more detail. *"Greedy"* quantifiers like *"*"* match as much as possible whereas *"lazy"* quantifiers such as *"*?"* match as little as possible to satisfy the search pattern of a composed regular expression (Friedl 2006, 141; Goyvaerts and Levithan 2012, 75-78).[4]

In addition, regular expressions can be modified in the way they are interpreted and applied using different regular expression modes (modifiers). These modifiers allow to change the search pattern of a particular regular expression (matching mode) in modern programming languages or in application programs. Regular expressions equipped with *"case-insensitive match mode"* ((?i)) ignore the letter case of the input (textual elements) during the matching process allowing the search pattern to match both upper and lower case letters (Friedl 2006, 110-111; Goyvaerts and Levithan 2012, 29). Since modern applications work with multiple (coding) lines regular expressions need to be modified in order to match a string across different lines. *"Dot-matches-all match mode"* also known as *"single-line mode"* ((?s)) modifies the search pattern of a regular expression in a way that it matches a character string across multiple lines (Friedl 2006, 110-113; Goyvaerts and Levithan 2012, 39-40). By designing regular

[4] *"Nondeterministic Finite Automation (NFA)"* or *"regex-directed"* and *"Deterministic Finite Automation (DFA)"* or *"text-directed"* engines are the two basic kinds of regular expression engines (Friedl 2006, 153-156). Lazy quantifiers can only be implemented in *"regex-directed"* regular expression engines (Goyvaerts 2007, 7). The DFA engine type is used by awk, egrep, flex, lex, MySQL and Procmail. The NFA engine type is used by GNU Emacs, Java, grep, less, more, .Net languages, PCRE library, Perl, PHP, Python, Ruby, sed and vi (Friedl 2006, 145).

expressions and implementing them into modern computer software the results of various search patterns (textual information) can be highlighted and changed or even removed from the underlying text at all (Friedl 2006, 2; Goyvaerts and Levithan 2012, 1).

5.2 Extraction of Textual Information

Researchers in the field of finance and accounting (as well as business data providers) use the *"Complete Submission Text Files"* (file extension *.txt) provided by the SEC and its EDGAR system to extract textual information from financial statements. In order to delete all non-textual elements (HTML tags and their corresponding attributes) most often special text-processing programs and their predefined applications (HTML-Parser) are used. This again is problematic for several reasons. First, using predefined text-processing operators to delete non-textual elements makes one platform-dependent since a specific HTML-Parser can not be (easily) implemented into any other text-processing program in use. Secondly, since the extraction algorithm of the HTML-Parser is complex or not presented at all its extraction results can hardly be validated. Thirdly, because of these drawbacks empirical research results are challenging to replicate for a particular or any other data sample. Regular expressions can in fact overcome these problems in extracting textual information embedded in financial statements filed with the SEC. They offer platform-independent (research) results which can be validated and replicated for any data sample at any given time.

The proposed extraction algorithm (*"Annual Report Algorithm"*) first decomposes the *"Complete Submission Text File"* (file extension *.txt) into its components (RegEx 1). In the end, the entire algorithm is validated through obtaining exactly one core (Form 10-K) document and the number of exhibits which have been embedded in the *"Complete Submission Text File"* for every financial statement in the data sample. Next, the *"Annual Report Algorithm"* identifies all other file types contained in the submission since these additional documents are not either a core document or an exhibit within the text version of the filing (RegEx 2). Table 6 illustrates the regular expressions needed to decompose the *"Complete Submission Text File"* of a financial statement filed with the SEC and to identify the embedded document (file) types.

Table 6. Regular expressions contained in the *"Annual Report Algorithm"*.

ID	Description	Regular Expression
1	Decomposition of *"Complete Submission Text File"*	(?s)<DOCUMENT>.*?</DOCUMENT>
2	Identification of document (file) types	<TYPE>.*

Notes: The table presents the regular expressions contained in the *"Annual Report Algorithm"* for extracting documents and identifying document (file) types.

In addition to the filing components described earlier (10-K section, exhibits, XBRL, graphics), several other document (file) types might be embedded in financial statements such as MS Excel files (file extension *.xlsx), ZIP files (file extension *.zip) and encoded PDF files (file extension *.pdf). By applying additional rules in the *"Annual Report Algorithm"* (RegExes 3-22) these documents are deleted to be able to extract textual information only from the core document and the various exhibits contained in the *"Complete Submission Text File"*. The additional SEC-header is not supposed to be removed separately since it has already been deleted by the algorithm.[5] Table 7 illustrates the regular expressions applied to delete document (file) types other than the core document and the corresponding exhibits.

Table 7. Regular expressions contained in the *"Annual Report Algorithm"*.

ID	Description	Regular Expression
3	Removal of graphic files	(?s)<TYPE>GRAPHIC.*?</TEXT>
4	Removal of MS Excel files	(?s)<TYPE>EXCEL.*?</TEXT>
5	Removal of PDF files	(?s)<TYPE>PDF.*?</TEXT>
6	Removal of ZIP files	(?s)<TYPE>ZIP.*?</TEXT>
7	Removal of cover letter	(?s)<TYPE>COVER.*?</TEXT>
8	Removal of correspondence between SEC staff and EDGAR participant	(?s)<TYPE>CORRESP.*?</TEXT>
9	Removal of XBRL instance document	(?s)<TYPE>EX-10[01].INS.*?</TEXT>
10	Removal of XBRL instance document	(?s)<TYPE>EX-99.SDR [KL].INS.*?</TEXT>
11	Removal of XBRL taxonomy extension schema document	(?s)<TYPE>EX-10[01].SCH.*?</TEXT>
12	Removal of XBRL taxonomy extension schema document	(?s)<TYPE>EX-99.SDR [KL].SCH.*?</TEXT>
13	Removal of XBRL taxonomy extension linkbase document	(?s)<TYPE>EX-10[01].CAL.*?</TEXT>
14	Removal of XBRL taxonomy extension linkbase document	(?s)<TYPE>EX-99.SDR [KL].CAL.*?</TEXT>
15	Removal of XBRL taxonomy extension definition linkbase document	(?s)<TYPE>EX-10[01].DEF.*?</TEXT>
16	Removal of XBRL taxonomy extension definition linkbase document	(?s)<TYPE>EX-99.SDR [KL].LAB.*?</TEXT>
17	Removal of XBRL taxonomy extension labels linkbase document	(?s)<TYPE>EX-10[01].LAB.*?</TEXT>
18	Removal of XBRL taxonomy extension labels linkbase document	(?s)<TYPE>EX-99.SDR [KL].LAB.*?</TEXT>
19	Removal of XBRL taxonomy extension presentation linkbase document	(?s)<TYPE>EX-10[01].PRE.*?</TEXT>
20	Removal of XBRL taxonomy extension presentation linkbase document	(?s)<TYPE>EX-99.SDR [KL].PRE.*?</TEXT>
21	Removal of XBRL taxonomy extension reference linkbase document	(?s)<TYPE>EX-10[01].REF.*?</TEXT>
22	Removal of XBRL documents	(?s)<TYPE>XML.*?</TEXT>

Notes: The table presents the regular expressions contained in the *"Annual Report Algorithm"* for deleting nonrelevant document (file) types.

Next, the *"Annual Report Algorithm"* deletes all metadata included in the core document and the exhibits (RegExes 23-27). Table 8 illustrates the regular expressions for deleting metadata in SEC EDGAR documents.

[5] RegEx 1 extracts filing parts declared as documents. The SEC header disclosed at the top of every *"Complete Submission Text File"* is declared as "<SEC-HEADER>", therefore the head section is not a separate document. The SEC header contains additional information about a particular filer such as the industry classification and the address. The phrases "BEGIN PRIVACY-ENHANCED MESSAGE" and "END PRIVACY-ENHANCED MESSAGE" in early SEC EDGAR complete submission text filings are not deleted separately for the same reason.

Table 8. Regular expressions contained in the *"Annual Report Algorithm"*.

ID	Description	Regular Expression
23	Removal of document type information	<TYPE>.*
24	Removal of sequence information	<SEQUENCE>.*
25	Removal of filename	<FILENAME>.*
26	Removal of description	<DESCRIPTION>.*
27	Removal of head section (including document title)	(?s)<HEAD>.*?</HEAD>

Notes: The table presents the regular expressions contained in the *"Annual Report Algorithm"* for deleting nonrelevant document metadata.

Before deleting all HTML elements and their corresponding attributes (RegEx 29) the algorithm deletes tables since they contain non-textual (quantitative) information (RegEx 28).[6] Table 9 illustrates the set of regular expressions applied to delete tables and HTML elements embedded in financial statements filed with the SEC.

Table 9. Regular expressions contained in the *"Annual Report Algorithm"*.

ID	Description	Regular Expression
28	Removal of table content	(?s)(?i)<Table.*?</Table>
29	Removal of HTML tags and attributes	(?s)<[^>]*>

Notes: The table presents the regular expressions contained in the *"Annual Report Algorithm"* for deleting tables and HTML elements.

After extracting the core document and the exhibits as well as deleting all HTML elements, the *"Annual Report Algorithm"* adjusts the content embedded in the body section of each HTML-formatted document in order to extract textual elements from financial statements on the EDGAR database. According to the SEC filer manual the EDGAR system suspends financial statements which contain extended ASCII characters. However, it supports submissions with extended character references. By using ISO-8859-1/Latin-1 decimal character references or entity-names (either technique is allowed within SEC submissions) extended ASCII characters can be embedded in financial statement submissions. These extended character sets within HTML documents included in the *"Complete Submission Text File"* need to be decoded to be able to extract human-readable textual information from financial statements (Filer Manual 2016, Section 5-19). The *"Annual Report Algorithm"* finally decodes all extended character sets (RegExes 30-680) most likely embedded in financial statements filed with the SEC and its EDGAR system formatted in HTML 4.01 (ASCII, ANSI/Windows-1252, ISO-8859-1/Latin-1, mathematical, Greek, symbolic and special characters).[7]

[6] SEC EDGAR filers might use HTML table tags to structure textual content in electronic submissions.

[7] Instead of applying RegExes 30-680 investors might want to use a predefined text processing operator since they are less dependent on extraction results that can be replicated for other data samples.

5.3 Extraction Results

By applying the *"Annual Report Algorithm"* investors and researchers are able to extract textual information from financial statements filed with the SEC for thousands of companies in a fully automated process. Based on the *"Complete Submission Text File"* provided by the EDGAR system the algorithm extracts the core (Form 10-K) document and the exhibits which have been embedded in the text version of a company´s financial statement. For example for Coca Cola´s 2015 annual report on Form 10-K filed on February 25, 2016 via EDGAR the algorithm extracts one core document in addition to ten different exhibits. Figure 4 illustrates partial extraction results for the 10-K section of the annual report as well as for two exhibits.

Figure 4. Examples of the extraction result of the *"Annual Report Algorithm"*.

UNITED STATES SECURITIES AND EXCHANGE COMMISSION Washington, D.C. 20549 **FORM 10-K** For the fiscal year ended December 31, 2015 OR For the transition period from to Commission File No. 001-02217 (Exact name of Registrant as specified in its charter) Registrant's telephone number, including area code: (404) 676-2121 Securities registered pursuant to Section 12(b) of the Act: Securities registered pursuant to Section 12(g) of the Act: None…

Exhibit 23.1 CONSENT OF INDEPENDENT REGISTERED PUBLIC ACCOUNTING FIRM We consent to the incorporation by reference in the registration statements and related prospectuses of The Coca-Cola Company listed below of our reports dated February 25, 2016, with respect to the consolidated financial statements of The Coca-Cola Company and subsidiaries, and the effectiveness of internal control over financial reporting of The Coca-Cola Company and subsidiaries, included in this Annual Report (Form 10-K) for the year ended December 31, 2015. /s/ ERNST & YOUNG LLP Atlanta, Georgia February 25, 2016…

EXHIBIT 31.1 CERTIFICATIONS I, Muhtar Kent, Chairman of the Board of Directors and Chief Executive Officer of The Coca-Cola Company, certify that: 1. I have reviewed this annual report on Form 10-K of The Coca-Cola Company; 2. Based on my knowledge, this report does not contain any untrue statement of a material fact or omit to state a material fact necessary to make the statements made, in light of the circumstances under which such statements were made, not misleading with respect to the period covered by this report…

Notes: The figure presents extraction results from Coca Cola´s 2015 annual report on Form 10-K filed with the SEC. The first part of the figure displays the actual 10-K section embedded in text version of the submission. The second part shows the statement of the auditing firm. The certification of the annual report by the CEO is presented in the last part of the figure.

Besides from textual content of entire documents (10-K section and exhibits) contained in the *"Complete Submission Text File"* investors and researchers might be interested in extracting textual information from particular sections (Items) within the core 10-K section of an annual report (like Item 1A - Risk Factors; Item 3 - Legal Proceedings; Item 7 - Management's Discussion and Analysis of Financial Condition and Results of Operations etc.). In order to extract textual information from particular 10-K items the *"Annual Report Algorithm"* is modified to the *"Items Algorithm"*. Excluding all exhibits, the modified *"Items Algorithm"* isolates only the 10-K section within the SEC submission. After deleting nonrelevant information and decoding reserved characters

within the document investors and researchers can extract textual information from specific 10-K items.[8] Table 10 specifies the modified *"Items Algorithm"* applied to extract textual information from particular items of the annual report on Form 10-K filed with the SEC.

Table 10. Regular expressions contained in the *"Items Algorithm"*.

ID	Description		Regular Expression
1.1	Extraction of 10-K section		(?s)<TYPE>10-K.*?</TEXT>
2.1	Removal of document metadata		RegExes 23-28
3.1	Removal of table content		(?s)(?i)<Table.*?</Table>
4.1	Decoding of reserved characters		See RegExes 30-680
5.1	Identification and renaming of item headings (">°Item")		(?s)(?i)(?m)> +Item|>Item|^Item
6.1	Removal of multiple empty spaces		(?s) +
7.1	Extraction of **Item 1.**	- Business	(?s)(?i)°Item 1[^AB012345].*?°Item
7.2	Extraction of **Item 1A.**	- Risk Factors	(?s)(?i)°Item 1A.*?°Item
7.3	Extraction of **Item 1B.**	- Unresolved Staff Comments	(?s)(?i)°Item 1B.*?°Item
7.4	Extraction of **Item 2.**	- Properties	(?s)(?i)°Item 2.*?°Item
7.5	Extraction of **Item 3.**	- Legal Proceedings	(?s)(?i)°Item 3.*?°Item
7.6	Extraction of **Item 4.**	- Mine Safety Disclosures	(?s)(?i)°Item 4.*?°Item
7.7	Extraction of **Item X.**	- Executive Officers of the Company	(?s)(?i)°Item X.*?°Item
7.8	Extraction of **Item 5.**	- Market for Registrant's Common Equity, Related Stockholder Matters and Issuer Purchases of Equity Securities	(?s)(?i)°Item 5.*?°Item
7.9	Extraction of **Item 6.**	- Selected Financial Data	(?s)(?i)°Item 6.*?°Item
7.10	Extraction of **Item 7.**	- Management's Discussion and Analysis of Financial Condition and Results of Operations	(?s)(?i)°Item 7[^A].*?°Item
7.11	Extraction of **Item 7A.**	- Quantitative and Qualitative Disclosures About Market Risk	(?s)(?i)°Item 7A.*?°Item
7.12	Extraction of **Item 8.**	- Financial Statements and Supplementary Data	(?s)(?i)°Item 8.*?°Item
7.13	Extraction of **Item 9.**	- Changes in and Disagreements with Accountants on Accounting and Financial Disclosure	(?s)(?i)°Item 9[^AB].*?°Item
7.14	Extraction of **Item 9A.**	- Controls and Procedures	(?s)(?i)°Item 9A.*?°Item
7.15	Extraction of **Item 9B.**	- Other Information	(?s)(?i)°Item 9B.*?°Item
7.16	Extraction of **Item 10.**	- Directors, Executive Officers and Corporate Governance	(?s)(?i)°Item 10.*?°Item
7.17	Extraction of **Item 11.**	- Executive Compensation	(?s)(?i)°Item 11.*?°Item
7.18	Extraction of **Item 12.**	- Security Ownership of Certain Beneficial Owners and Management and Related Stockholder Matters	(?s)(?i)°Item 12.*?°Item
7.19	Extraction of **Item 13.**	- Certain Relationships and Related Transactions, and Director Independence	(?s)(?i)°Item 13.*?°Item
7.20	Extraction of **Item 14.**	- Principal Accounting Fees and Services	(?s)(?i)°Item 14.*?(°Item|</TEXT>)
7.21	Extraction of **Item 15.**	- Exhibits, Financial Statement Schedules	(?s)(?i)°Item 15[^°]*?</TEXT>
8.1	Removal of HTML tags and attributes		(?s)<[^>]*>

Notes: The table presents the regular expressions contained in the modified *"Items Algorithm"* for extracting particular items from the annual report on Form 10-K. RegExes 1.1-6.1 modify the text version of a financial statement to be able to extract (clear) textual information from particular items. RegExes 7.1-7.21 represent the actual regular expressions designed to extract particular sections from the text version of the annual report.

[8] HTML-formatting inconsistencies might influence the capabilities of the *"Items Algorithm"* in extracting and separating Form 10-K items from multiple filings. Parsing errors due to formatting inconsistencies across documents (SEC filings) have the potential to produce "extraordinary results" in form of misspecified 10-K sections (Loughran and McDonald 2016, 1218). Despite several SEC EDGAR filers using HTML table tags to structure textual content (section headings and section content) tables are deleted due to a majority of research studies in the field of textual analysis excluding table content in general. By deleting tables just before removing HTML tags (RegEx 8.1) the capability of the *"Items Algorithm"* in extracting certain items from the entire 10-K section can be enhanced. At the same time forfeiting the ability to quantify this capability in extracting different sections in relation to the overall text length (several sections would not appear in the overall text length of the entire 10-K section due to being embedded in table tags).

Using only regular expressions to extract textual information from financial statements investors and researchers can implement the designed extraction algorithms in any modern application and computer program available today. By applying either the *"Annual Report Algorithm"* or the *"Items Algorithm"* entire documents (10-K section and exhibits) or particular items from the core 10-K section can be extracted from the annual SEC submissions in order to be analyzed. More importantly, while compensating for expensive commercial products the algorithms and their extraction results can be validated and replicated for any data sample at any given time. Figure 5 finally illustrates several extraction results of the *"Items Algorithm"* from the annual report on Form 10-K highly relevant to investors and researchers alike.

Figure 5. Examples of the extraction result of the *"Items Algorithm"*.

°**Item 1A. RISK FACTORS** In addition to the other information set forth in this report, you should carefully consider the following factors, which could materially affect our business, financial condition or results of operations in future periods. The risks described below are not the only risks facing our Company. Additional risks not currently known to us or that we currently deem to be immaterial also may materially adversely affect our business, financial condition or results of operations in future periods…

°**Item 3. LEGAL PROCEEDINGS** The Company is involved in various legal proceedings, including the proceedings specifically discussed below. Management believes that the total liabilities to the Company that may arise as a result of currently pending legal proceedings will not have a material adverse effect on the Company taken as a whole. Aqua-Chem Litigation On December 20, 2002, the Company filed a lawsuit (The Coca-Cola Company v. Aqua-Chem, Inc., Civil Action No. 2002CV631-50) in the Superior Court of Fulton County, Georgia…

°**Item 7. MANAGEMENT'S DISCUSSION AND ANALYSIS OF FINANCIAL CONDITION AND RESULTS OF OPERATIONS** Overview The following Management's Discussion and Analysis of Financial Condition and Results of Operations ("MD&A") is intended to help the reader understand The Coca-Cola Company, our operations and our present business environment. MD&A is provided as a supplement to - and should be read in conjunction with - our consolidated financial statements and the accompanying notes thereto contained in "Item 8. Financial Statements and Supplementary Data" of this report. This overview summarizes the MD&A, which includes the following sections…

Notes: The figure presents extraction results from Coca Cola´s 2015 annual report on Form 10-K filed with the SEC. The first part of the figure displays Item 1A (Risk Factors) embedded in the overall 10-K section. The last two parts of the figure show Item 3 (Legal Proceedings) and Item 7 (Management´s Discussion and Analysis of Financial Condition and Results of Operations) contained in the 10-K section of the *"Complete Submission Text File"*.

6. Validation of Extraction Algorithms

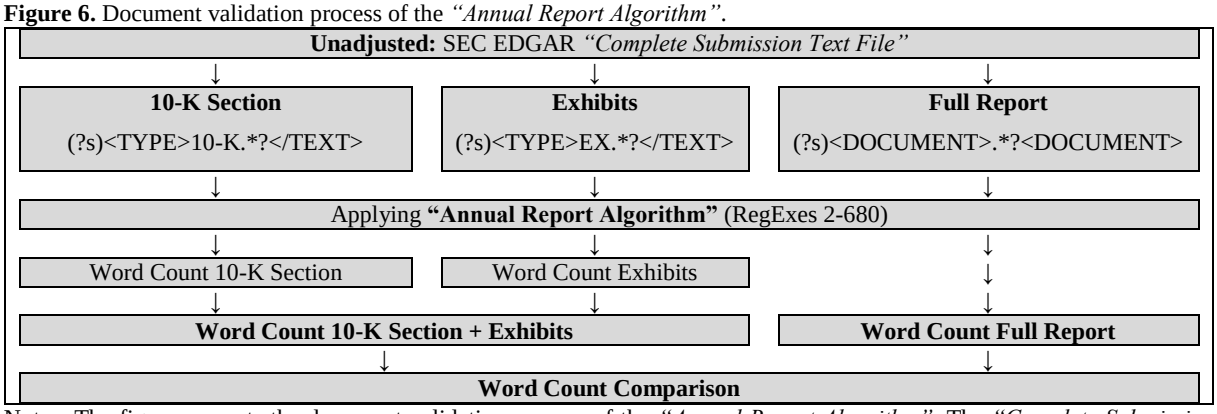
In order to validate the proposed extraction algorithms and to test their capabilities, I retrieve all Form 10-K filings listed in the SEC EDGAR form index files. Using the data gathering method as described in Section 3 in total 188,875 annual reports (167,599 on Form 10-K and 21,276 on Form 10-K405[9]) filed between 1993 and 2016 are retrieved from the EDGAR database (SEC EDGAR Form 10-K types as used in Loughran and McDonald 2011a).[10] The *"Annual Report Algorithm"* is applied to all submissions to derive different word counts for each filing made

[9] Before eliminated by the SEC in 2003 due to confusion and inconsistency in its application an annual report on Form 10-K405 (405 classification) was used to indicate that a disclosure of delinquent filers pursuant to Item 405 was not included in the current 10-K filing (Loughran and McDonald 2011a, 39).

[10] One out of 188,876 annual reports (Form: 10-K405/ CIK: 884219/ Year: 1996) could not be retrieved (https://www.sec.gov/Archives/edgar/data/884219/0000884219-96-000033.txt).

with the SEC. In addition to the overall word count of an annual report, for each core document (10-K section) and the exhibits embedded in a *"Complete Submission Text File"* an individual word count is retrieved in order to be compared (XBRL files declared as exhibits are deleted). Figure 6 illustrates how word counts for each filing and its components are obtained from the *"Complete Submission Text File"* for the document validation process of the *"Annual Report Algorithm"*.

Figure 6. Document validation process of the *"Annual Report Algorithm"*.



Notes: The figure presents the document validation process of the *"Annual Report Algorithm"*. The *"Complete Submission Text File"* of each financial statement as provided on the SEC server is used to extract all relevant components (documents). The *"Annual Report Algorithm"* is applied to each filing in order to retrieve word counts for all relevant documents embedded in the submission. The word count of all relevant documents is compared with the overall length of the submission. A mismatch between the word counts would indicate that the entire report contains nonrelevant document (file) types after applying the *"Annual Report Algorithm"*.

This word count comparison between the overall report on full length and its different components cannot be a validation of the *"Annual Report Algorithm"* since the same algorithm is simply applied to different sets of textual information (10-K section, exhibits, full report). However, if the entire report would still contain document (file) types or elements which are not a part of the core 10-K section or a corresponding exhibit the word count of a certain financial statement would be artificially increased (Word Count Full Report). Table 11 presents the word count comparison of all annual reports and their embedded documents.

Table 11. Results of the document validation process.

Year	Filings	Word Count Comparison				
		10-K Sections + Exhibits	Full Reports	Extraction Error		
				%	Filings	Type
2016	6,467	298,196,698	298,196,698	0.0000	0	-
2015	7,985	350,610,193	350,610,193	0.0000	0	-
2014	8,084	351,653,600	351,664,987	0.0032	1	- Included Form 10-Q
2013	8,105	355,676,863	355,676,863	0.0000	0	-
2012	8,393	347,050,833	347,082,370	0.0091	2	- Declaration error - Included Form 10-Q
2011	8,840	363,208,197	363,208,197	0.0000	0	-
2010	9,165	371,951,148	371,951,148	0.0000	0	-
2009	9,839	397,555,819	397,555,819	0.0000	0	-
2008	8,746	342,695,246	342,695,246	0.0000	0	-
2007	8,574	340,907,681	340,907,681	0.0000	0	-
2006	8,852	326,730,970	326,730,970	0.0000	0	-
2005	9,017	326,083,277	326,111,171	0.0086	1	- Declaration error
2004	8,567	330,972,848	330,972,848	0.0000	0	-
2003	8,468	331,839,061	331,889,877	0.0153	2	- Declaration errors
2002	8,927	332,483,818	332,574,108	0.0272	14	- Declaration errors
2001	9,248	325,089,007	325,099,248	0.0032	1	- Declaration error
2000	9,869	335,180,436	335,242,190	0.0184	5	- Declaration errors - Syntax error
1999	10,122	340,328,009	340,444,185	0.0341	10	- Declaration errors - Syntax errors
1998	10,287	363,400,331	363,480,640	0.0221	8	- Declaration errors - Syntax errors
1997	9,899	319,367,142	319,428,271	0.0191	5	- Declaration errors - Syntax errors
1996	6,258	181,885,167	181,911,885	0.0147	6	- Declaration errors - Syntax errors
1995	3,236	112,604,017	112,622,922	0.0168	3	- Declaration errors - Syntax error
1994	1,923	76,399,982	76,399,982	0.0000	0	-
1993	4	82,283	82,283	0.0000	0	-
Total	**188,875**	**7,221,952,626**	**7,222,539,782**	**0.0081**	**58**	**-**

Notes: The table presents the document validation results of the *"Annual Report Algorithm"*. The third column displays the total word count of all reports in a particular year if only relevant sections have been included. The next column shows the actual word count of all reports in a particular year retrieved from the *"Complete Submission Text Files"* by applying the *"Annual Report Algorithm"*. Mismatches between the word counts for every year are shown in the next column indicating that the algorithm is not able to delete all nonrelevant document (file) types contained in the submissions.

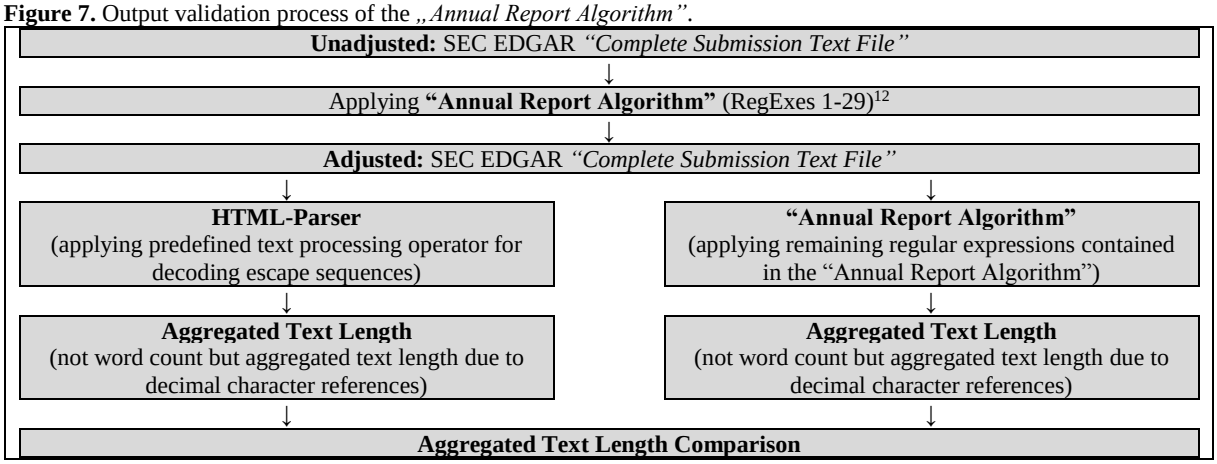
In fact, the ability to validate the entire extraction procedure by applying an alternative to the *"Annual Report Algorithm"* (e.g. HTML-Parser) is limited since to a certain extent the same regular expressions have to be used to create the input for both extraction methods in the first place (extracting core 10-K document and exhibits, deleting nonrelevant document (file) types etc.). Due to this disability in validating the entire extraction process from the beginning by applying an HTML-Parser one has to validate the input the proposed algorithm is creating and its extraction results separately, therefore validating the entire information extraction process. The validation of the textual input created by the *"Annual Report Algorithm"* is represented by the extraction algorithm itself since it uses only regular expressions combined with the electronic filing requirements introduced by the SEC

(precisely not the SEC but Attain, LLC)[11]. According to the SEC, all documents embedded in a *"Complete Submission Text File"* must be equipped with a "<TYPE>" tag representing the conformed document type of that particular submission part within the text version of the filing (<TYPE>10-K, <TYPE>10-Q, <TYPE>8-K, <TYPE>EX-1, <TYPE>EX-2 etc.) (SEC EDGAR 2015, 12-13, 29). The *"Annual Report Algorithm"* (RegExes 1-29) uses these requirements in order to extract the core document and the corresponding exhibits from annual reports while deleting all documents associated with XBRL and other document (file) types. The search patterns of the *"Annual Report Algorithm"* which have been designed accordingly to the filing requirements of the SEC can be validated due to the general pattern notation of the regular expression language.

An output comparison between the *"Annual Report Algorithm"* and a common HTML-Parser shall serve as an additional validation for the remaining extraction procedure. Therefore, I modify the *"Complete Submission Text Files"* as provided by the SEC (unadjusted filings) and apply the first part of the *"Annual Report Algorithm"* (RegExes 1-29) in order to make the text version of the financial statements readable for the predefined HTML-Parser (adjusted filings). Since this part of the overall validation process focuses on how well the *"Annual Report Algorithm"* is capable of decoding escape sequences embedded in a *"Complete Submission Text File"* the aggregated text length of both procedures are compared rather than the word counts due to decimal character encodings (a simple word count comparison would not fully capture the disability of the *"Annual Report Algorithm"* in decoding these character references in relation to the HTML-Parser). Figure 7 illustrates the output validation process of the *"Annual Report Algorithm"*.

[11] The EDGAR Public Dissemination Service (PDS) System is a privatized PDS System managed by Attain, LLC. The system is the primary source to receive all accepted and valid EDGAR filings. The system became operational on November 1, 1998 (SEC EDGAR 2016).

Figure 7. Output validation process of the *„Annual Report Algorithm"*.



Notes: The figure presents the output validation process of the *"Annual Report Algorithm"*. The *"Complete Submission Text File"* of each financial statement as provided on the SEC server is adjusted in order to compare the output of the algorithm with the output a common HTML-Parser would produce. RegExes 1-29 modify the unadjusted document as provided on the EDGAR database before applying a predefined text processing operator (HTML-Parser). The aggregated text length for all filings of both procedures is compared in order to validate the capability of the *„Annual Report Algorithm"* in decoding escape sequences. The aggregated text length includes each individual element in an underlying text document (text, digits, spaces, special characters etc.).

Table 12 presents the validation results for the *"Annual Report Algorithm"*.

[12] HTML tags and their corresponding attributes are removed (RegEx 29) before applying the HTML-Parser in order to decode HTML escape sequences since unescaping might interfere HTML structure.

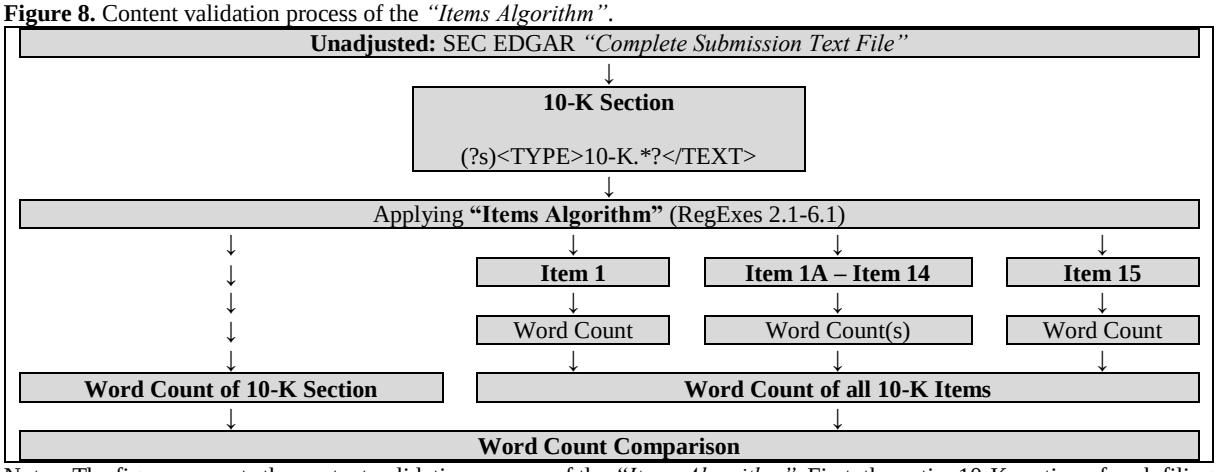
Table 12. Validation results of the *"Annual Report Algorithm"*.

Year	Filings	"Annual Report Algorithm"					
		No Parsing	HTML-Parser	Annual Report Algorithm	Error (%)		
					Mean	Med.	Max.
2016	6,467	101.84	100.00	100.08	0.08	0.01	5.93
2015	7,985	102.04	100.00	100.06	0.06	0.01	5.42
2014	8,084	102.14	100.00	100.05	0.05	0.01	3.54
2013	8,105	102.14	100.00	100.04	0.05	0.00	6.10
2012	8,393	102.40	100.00	100.05	0.05	0.00	9.62
2011	8,840	102.63	100.00	100.05	0.05	0.01	19.07
2010	9,165	102.77	100.00	100.04	0.05	0.01	8.80
2009	9,839	102.92	100.00	100.04	0.04	0.00	2.89
2008	8,746	102.95	100.00	100.05	0.05	0.00	5.39
2007	8,574	102.41	100.00	100.04	0.05	0.00	2.72
2006	8,852	102.59	100.00	100.04	0.05	0.00	3.02
2005	9,017	102.80	100.00	100.04	0.04	0.00	2.87
2004	8,567	102.49	100.00	100.05	0.05	0.00	40.63
2003	8,468	102.28	100.00	100.03	0.04	0.00	2.23
2002	8,927	101.46	100.00	100.02	0.02	0.00	5.30
2001	9,248	100.53	100.00	100.01	0.01	0.00	4.34
2000	9,869	100.17	100.00	100.01	0.01	0.00	4.67
1999	10,122	100.01	100.00	100.00	0.00	0.00	1.08
1998	10,287	100.00	100.00	100.00	0.00	0.00	0.03
1997	9,899	100.00	100.00	100.00	0.00	0.00	0.16
1996	6,258	100.00	100.00	100.00	0.00	0.00	0.90
1995	3,236	100.00	100.00	100.00	0.00	0.00	0.00
1994	1,923	100.00	100.00	100.00	0.00	0.00	0.00
1993	4	100.00	100.00	100.00	0.00	0.00	0.00
Total	**188,875**	**101.95**	**100.00**	**100.04**	**0.03**	**0.00**	**40.63**

Notes: The table presents the validation results of the *"Annual Report Algorithm"*. The third column presents the aggregated text length including all elements in a filing (text, digits, spaces, special characters etc.) as provided on the SEC server for all submissions in a particular year. The next column represents the aggregated text length of all submissions in a certain year if an HTML-Parser is used to decode character references within the submissions. The fifth column shows the aggregated text length of all filings when using the *"Annual Report Algorithm"* instead of a predefined text processing operator (HTML-Parser) to decode the embedded character references. The inaccuracy in decoding escape sequences when using the *"Annual Report Algorithm"* rather than applying an HTML-Parser is shown in the next column.

In contrast to the *"Annual Report Algorithm"* the modified *"Items Algorithm"* is validated by its ability to distribute the extracted information to the individual items an annual report filed with the SEC is composed of. In order to test and validate the capabilities of the *"Items Algorithm"* I again use the *"Complete Submission Text Files"* as provided by the SEC and extract only the 10-K section of each filing. For each submission, I retrieve separate word counts for the 10-K section and for all individual items extracted by the *"Items Algorithm"*. Despite textual information embedded in the 10-K section not contained in a particular item (introduction) a word count comparison between the overall 10-K section and all items represents an attempt to validate the capabilities of the *"Items Algorithm"* in extracting certain sections from the core document of an annual report filed with the SEC and its EDGAR system. Figure 8 illustrates the content validation process of the *"Items Algorithm"*.

Figure 8. Content validation process of the *"Items Algorithm"*.



Notes: The figure presents the content validation process of the *"Items Algorithm"*. First, the entire 10-K section of each filing from the *"Complete Submission Text File"* as provided on the SEC server is extracted. Word counts for the entire 10-K section as well as for all individual items are retrieved by applying the *"Items Algorithm"* in order to be compared. Due to structural changes of the annual report on Form 10-K over time (different number of items) the relation of text length between the overall 10-K section and all individual items shall represent the ability of the algorithm to extract particular items from the 10-K section.

Table 13 presents the validation results for the *"Items Algorithm"*.

Table 13. Validation results of the *"Items Algorithm"*.

Year	Filings		"Items Algorithm"								
			Word Count Comparison		Precision, Recall, and F-measure						
					Filings	Items					
	Number	%	∑ of Items (%)	Rest/ Error (%)	Tested	Exists	Extracted	Correct	Precision	Recall	F-measure
2016	2,886	44.63	97.72	2.28	10	195	191	185	96.86	94.87	95.85
2015	3,714	46.51	97.60	2.40	10	200	200	195	97.50	97.50	97.50
2014	4,004	49.53	97.52	2.48	10	198	197	193	97.97	97.47	97.72
2013	3,962	48.88	97.53	2.47	10	192	188	188	100.00	97.92	98.95
2012	3,938	46.92	97.39	2.61	10	198	192	183	95.31	92.42	93.85
2011	4,104	46.43	97.43	2.57	10	193	191	189	98.95	97.93	98.44
2010	2,805	30.61	97.10	2.90	10	197	168	155	92.26	78.68	84.93
2009	2,719	27.63	97.15	2.85	10	196	181	164	90.61	83.67	87.00
2008	2,077	23.75	97.28	2.72	10	196	184	170	92.39	86.73	89.47
2007	2,065	24.08	97.38	2.62	10	195	184	173	94.02	88.72	91.29
2006	2,662	30.07	97.49	2.51	10	198	184	165	89.67	83.33	86.39
2005	3,122	34.62	97.70	2.30	10	181	175	163	93.14	90.06	91.57
2004	3,496	40.81	97.60	2.40	10	173	170	163	95.88	94.22	95.04
2003	3,903	46.09	97.33	2.67	10	161	161	154	95.65	95.65	95.65
2002	4,961	55.57	97.65	2.35	10	150	150	150	100.00	100.00	100.00
2001	5,799	62.71	97.61	2.39	10	146	144	135	93.75	92.47	93.10
2000	6,268	63.51	97.55	2.45	10	150	149	138	92.62	92.00	92.31
1999	6,302	62.26	97.55	2.45	10	146	145	143	98.62	97.95	98.28
1998	6,492	63.11	97.56	2.44	10	140	140	128	91.43	91.43	91.43
1997	6,397	64.62	97.43	2.57	10	132	129	125	96.90	94.70	95.79
1996	3,918	62.61	97.27	2.73	10	136	121	112	92.56	82.35	87.16
1995	1,907	58.93	97.07	2.93	10	135	135	127	94.07	94.07	94.07
1994	1,039	54.03	97.23	2.77	10	140	139	133	95.68	95.00	95.34
1993	1	25.00	98.49	1.51	1	14	14	14	100.00	100.00	100.00
Total	**88,541**	**46.88**	**97.48**	**2.52**	231	3,962	3,832	3,645	95.12	92.00	93.53

Notes: The table presents the validation results of the *"Items Algorithm"*. The second and third columns show the number of filings of which items could be extracted from by applying the *"Items Algorithm"* (filings were not machine-parsable due to lacks of content, inconsistent filing structure, table tags and HTML formatting inconsistencies). Only filings with extracted items length exceeding 90 percent of 10-K section are presented. The next two columns show the average amount of extracted information from each filing in a particular year since 1993. The next columns show the performance evaluation of the *"Items Algorithm"* using precision (=number of correct answers/number of total answers), recall (=number of correct answers/total possible correct answers), and F-measure (=2*precision*recall/precision+recall).

7. Descriptive Statistics on Form 10-K contents

In total, I examine the textual composition of 188,875 annual reports filed with the SEC between 1993 and 2016. On average, an annual report on Form 10-K submitted to the EDGAR system during the sample period is composed of 38,240 words. The average word count of an annual submission increased from 39,730 in 1994 to 46,111 in 2016. The medians of the word counts increased accordingly. The majority of textual information embedded in an annual report on Form 10-K are contained in the core document (64.95 percent) whereas the disclosed exhibits represent only a minority of the overall textual elements stated in annual submissions (35.04 percent). By examining the EDGAR database and its Form 10-K filings in more detail, investors and researchers can see that the average file size (Megabyte) of an annual report made with the electronic disclosure system increased in recent years due to HTML formatting, ASCII-encodings and XBRL documents. Table 14 presents descriptive statistics of the text length and the file size of 188,875 annual reports on Form 10-K (Form 10-K405) filed with the SEC between 1993 and 2016.

Table 14. Descriptive statistics of SEC EDGAR Form 10-K reports.

Year	Filings	Word Count					File Size		
		Full Report (Number)			10-K Sections (%)	Exhibits (%)	Mean (MB)	Med. (MB)	Max. (MB)
		Mean	Med.	Max.					
2016	6,467	46,111	39,997	1,112,167	79.54	20.46	12.50	9.11	261.90
2015	7,985	43,909	37,262	1,657,009	79.51	20.49	15.12	10.18	414.52
2014	8,084	43,501	35,840	2,884,474	78.38	21.62	14.08	9.72	402.86
2013	8,105	43,884	35,181	6,257,121	77.32	22.68	13.22	9.38	254.18
2012	8,393	41,354	34,135	1,441,676	78.62	21.37	8.68	4.90	139.48
2011	8,840	41,087	33,008	1,031,964	77.33	22.67	4.48	1.71	212.57
2010	9,165	40,584	32,448	957,870	77.65	22.35	2.50	1.49	95.27
2009	9,839	40,406	32,074	3,997,528	74.97	25.03	1.90	1.33	86.21
2008	8,746	39,183	32,501	779,558	72.72	27.28	1.72	1.27	61.97
2007	8,574	39,761	32,206	2,617,579	73.67	26.33	1.81	1.28	91.99
2006	8,852	36,910	30,247	908,916	70.76	29.24	1.42	1.01	61.16
2005	9,017	36,166	28,854	1,442,810	66.13	33.86	1.19	0.82	80.62
2004	8,567	38,633	28,655	1,008,146	60.55	39.45	0.98	0.67	27.82
2003	8,468	39,193	28,738	911,982	58.15	41.83	0.90	0.55	24.01
2002	8,927	37,255	26,201	1,545,636	52.82	47.15	0.59	0.34	26.59
2001	9,248	35,153	24,531	1,308,749	52.03	47.97	0.40	0.28	23.34
2000	9,869	33,969	23,619	1,258,064	51.01	48.97	0.35	0.26	19.91
1999	10,122	33,634	23,290	496,458	49.40	50.56	0.33	0.25	8.29
1998	10,287	35,334	22,206	667,721	44.27	55.71	0.33	0.24	4.82
1997	9,899	32,269	20,496	650,347	44.84	55.14	0.30	0.22	4.82
1996	6,258	29,069	19,082	447,469	45.68	54.31	0.28	0.21	4.25
1995	3,236	34,803	22,570	361,832	38.51	61.48	0.34	0.24	4.03
1994	1,923	39,730	25,510	553,782	37.55	62.45	0.39	0.28	4.27
1993	4	20,571	18,247	31,993	83.01	16.99	0.23	0.26	0.27
Total	**188,875**	**38,240**	**28,772**	**6,257,121**	**64.95**	**35.04**	**3.56**	**0.71**	**414.52**

Notes: The table presents descriptive statistics of the text lengths, document compositions and file sizes for all annual reports filed with the SEC since 1993. Columns 3-5 show the means, medians and maxima of word counts of Form 10-K filings made on EDGAR. The average distribution of textual information between the 10-K sections and exhibits contained in the *"Complete Submission Text Files"* is presented in column 6 and 7. The last three columns of the table present the means, medians and maxima of the file sizes (Megabyte) for all Form 10-K submissions.

The distribution of textual elements among the various 10-K items is unequal. On average 22.65 percent of all textual information are contained in Item 1 (*"Business"*). Describing a company´s business as well as its main products and services, the item may also include information about the competition, regulations and other issues a particular company is faced with (SEC 2011; SEC Regulation S-K 2016, Section 229.101 Item 101). Item 7 (*"Management´s Discussion and Analysis of Financial Condition and Results of Operations – MD&A"*) represents 18.58 percent of the given information within Form 10-K filings made with the SEC. The item states information about a company´s operations and financial results in addition to its liquidity and capital resources. The section may include off-balance sheet arrangements and contractual obligations alongside key business risks (SEC 2011; SEC Regulation S-K 2016, Section 229.303 Item 303). Item 8 (*"Financial Statements and Supplementary Data"*) requires a company to disclose audited financial statements (SEC 2011; SEC Regulation S-X 2016, Section 210.3.01-210.3.02; SEC Regulation S-K 2016, Section 229.302 Item 302). Additional information explaining the financial statements in more detail (*"Notes to Consolidated Financial Statements"*, *"Report of Management"*, *"Report of Independent Registered Accounting Firm"* etc.) represent 15.96 percent of all given information in the

10-K section of an annual report. Item 1A (*"Risk Factors"*) describes significant factors that may adversely affect a filer´s business, financial condition or future financial performance (SEC 2011; SEC Regulation S-K 2016, Section 229.503 Item 503(c)). Since electronic filings became available on average 8.42 percent of all textual information disclosed in annual submissions are contained in this section. Each of the remaining items only represent a fraction of the overall textual information embedded in Form 10-K filings. While the length for most sections in annual reports remained constant over time the amount of textual information contained in Item 1A (*"Risk Factors"*) increased from 12.56 percent in 2006 to 20.10 percent in 2016 indicating that SEC EDGAR filers disclose more information about risks in recent years. Table 15 reports descriptive statistics of the distribution of textual information in Form 10-K reports.

Table 15. Distribution of textual information in SEC EDGAR Form 10-K reports.

Panel A

Year	Filings	10-K Sections	Item										
			1	1A	1B	2	3	4	X	5	6	7	7A
2016	2,886	100.00	16.71	20.10	0.03	0.74	1.08	0.36	0.00	1.45	1.35	17.82	1.65
2015	3,714	100.00	16.69	18.81	0.04	0.88	0.78	0.40	0.00	1.50	1.42	17.79	1.63
2014	4,004	100.00	17.22	17.40	0.08	0.89	0.92	0.40	0.00	1.53	1.47	17.96	1.60
2013	3,962	100.00	17.28	16.46	0.07	0.95	1.08	0.57	0.00	1.64	1.52	17.87	1.37
2012	3,938	100.00	17.58	15.45	0.07	1.00	0.98	0.53	0.00	1.67	1.53	17.88	1.66
2011	4,104	100.00	17.21	14.54	0.08	1.15	2.10	0.54	0.00	1.67	1.60	17.65	1.34
2010	2,805	100.00	17.53	14.35	0.11	1.03	1.08	0.87	0.00	1.70	1.69	17.82	1.33
2009	2,719	100.00	17.51	14.08	0.03	1.03	1.09	0.82	0.00	1.65	1.81	17.93	1.36
2008	2,077	100.00	18.16	13.95	0.06	1.03	1.31	1.14	0.00	1.27	1.91	18.45	1.56
2007	2,065	100.00	17.69	12.91	0.05	1.11	1.37	2.11	0.00	1.12	2.40	18.46	1.78
2006	2,662	100.00	19.05	12.56	0.22	1.09	1.33	1.26	0.01	1.03	2.04	19.76	1.71
2005	3,122	100.00	24.08	0.30	0.00	1.22	1.59	0.94	0.00	1.25	2.00	24.14	1.93
2004	3,496	100.00	24.33	0.03	0.01	1.50	1.73	1.36	0.00	1.26	2.14	23.48	1.95
2003	3,903	100.00	25.22	0.07	0.00	1.25	1.69	0.83	0.01	1.26	2.02	21.65	1.83
2002	4,961	100.00	26.97	0.07	0.00	1.37	1.62	0.73	0.01	1.34	1.40	20.93	1.96
2001	5,799	100.00	29.78	0.06	0.00	1.41	1.47	0.73	0.01	1.38	1.14	18.74	1.69
2000	6,268	100.00	29.91	0.04	0.00	1.64	1.48	1.02	0.01	1.47	1.10	17.94	1.38
1999	6,302	100.00	30.16	0.04	0.00	1.72	1.51	0.80	0.01	1.40	1.10	19.23	1.24
1998	6,492	100.00	31.27	0.03	0.00	1.80	1.59	0.99	0.01	1.49	1.13	16.87	0.37
1997	6,397	100.00	31.58	0.04	0.00	2.09	1.64	1.00	0.01	1.44	1.10	15.90	0.03
1996	3,918	100.00	29.80	0.03	0.01	2.27	2.24	1.06	0.01	1.45	0.89	15.31	0.00
1995	1,907	100.00	28.77	0.01	0.00	2.23	2.75	1.39	0.01	1.20	0.80	14.07	0.00
1994	1,039	100.00	27.78	0.02	0.00	2.26	3.97	1.54	0.01	1.10	0.72	12.61	0.00
1993	1	100.00	18.45	0.00	0.00	0.92	0.32	1.47	0.00	0.58	0.06	12.96	0.00
Total	88,541	100.00	22.65	8.42	0.04	1.29	1.42	0.83	0.01	1.44	1.51	18.58	1.39

Panel B

Year	Filings	10-K Sections	Item										Rest/ Error
			8	9	9A	9B	10	11	12	13	14	15	
2016	2,886	100.00	18.58	0.12	1.65	0.26	1.80	1.25	0.38	0.62	0.58	11.18	2.28
2015	3,714	100.00	18.38	0.15	1.80	0.26	2.20	1.58	0.50	0.75	0.65	11.41	2.40
2014	4,004	100.00	17.74	0.13	1.89	0.26	2.40	1.78	0.52	0.79	0.72	11.80	2.48
2013	3,962	100.00	17.26	0.15	2.00	0.25	2.59	1.94	0.57	0.86	0.80	12.29	2.47
2012	3,938	100.00	17.29	0.15	2.09	0.25	2.70	2.04	0.65	0.87	0.80	12.20	2.61
2011	4,104	100.00	16.35	0.19	2.16	0.27	2.89	2.70	0.68	0.95	0.95	12.42	2.57
2010	2,805	100.00	17.21	0.24	2.19	0.31	2.63	1.99	0.66	0.93	1.11	12.31	2.90
2009	2,719	100.00	17.21	0.22	2.31	0.24	2.29	1.79	0.72	0.79	0.91	13.36	2.85
2008	2,077	100.00	18.51	0.18	2.19	0.29	1.71	1.88	0.52	0.62	0.73	11.82	2.72
2007	2,065	100.00	17.33	0.15	1.83	0.34	1.44	2.33	0.50	0.53	0.79	13.13	2.62
2006	2,662	100.00	15.69	0.16	1.97	0.37	1.59	1.25	0.50	0.61	0.68	14.62	2.51
2005	3,122	100.00	17.18	0.24	1.79	0.41	1.73	1.31	0.58	0.64	0.85	15.53	2.30
2004	3,496	100.00	17.37	0.21	1.00	0.01	1.89	1.29	0.67	0.71	1.09	15.57	2.40
2003	3,903	100.00	16.46	0.79	0.11	0.00	1.70	1.43	0.74	0.80	1.80	17.66	2.67
2002	4,961	100.00	14.66	0.61	0.00	0.00	1.79	1.67	0.73	0.98	19.55	1.23	2.35
2001	5,799	100.00	14.17	0.53	0.00	0.00	2.03	1.78	0.81	1.02	20.85	0.02	2.39
2000	6,268	100.00	13.13	0.69	0.00	0.00	2.27	2.00	0.89	1.15	21.40	0.02	2.45
1999	6,302	100.00	12.50	0.72	0.00	0.00	2.33	1.98	0.88	1.12	20.77	0.04	2.45
1998	6,492	100.00	12.57	0.85	0.00	0.00	2.33	2.01	0.88	1.25	22.09	0.02	2.44
1997	6,397	100.00	12.71	0.99	0.00	0.00	2.54	1.94	0.87	1.18	22.35	0.02	2.57
1996	3,918	100.00	13.08	0.96	0.00	0.00	2.72	2.00	0.96	1.16	23.26	0.07	2.73
1995	1,907	100.00	14.87	0.89	0.00	0.00	2.36	1.67	0.74	0.81	24.46	0.03	2.93
1994	1,039	100.00	15.53	1.31	0.00	0.00	1.91	1.84	0.76	0.69	25.17	0.00	2.77
1993	1	100.00	0.09	0.09	0.00	0.00	0.29	0.13	0.27	0.16	62.70	0.00	1.51
Total	88,541	100.00	15.96	0.40	1.16	0.16	2.20	1.81	0.67	0.88	8.12	8.53	2.52

Notes: The table reports the average distribution of textual information among different 10-K items in annual reports filed with the SEC and its EDGAR database since 1993 (only filings with extracted items length exceeding 90 percent of 10-K section are presented).

8. Summary

This paper displays the huge amount and variety of publicly available corporate information filed with the SEC and distributed by its EDGAR database. It shows how massive data can be retrieved from the SEC server in a fast and efficient way using simple and easy accessible software. The second main purpose of this paper is to create standardized procedures (*"Annual Report Algorithm"* and *"Items Algorithm"*) investors and researchers can use to extract any kind of textual information from financial statements filed with the SEC. This is achieved by providing regular expressions for multiple steps of data cleaning and filtering. Using these dynamic and platform-independent extraction algorithms the paper analyses the textual composition of more than 180,000 annual reports filed with the SEC via the EDGAR system between 1993 and 2016. The algorithms are tested for validity in several ways. The tools and algorithms intend to reduce costs and lower technical boundaries for researchers in the field of finance and accounting to engage in textual analysis.

References

Bodnaruk A, Loughran T, McDonald B (2015) Using 10-K Text to Gauge Financial Constraints, in: Journal of Financial and Quantitative Analysis (JFQA), 50(4)/2015, 623-646

Bovee M, Kogan A, Nelson K, Srivastava R P, Vasarhelyi M A, (2005) Financial Reporting and Auditing Agent with Net Knowledge (FRAANK) and eXtensible Business Reporting Language (XBRL), in: Journal of Information Systems (JIS), 19(1)/2005, 19-41

Chakraborty V, Vasarhelyi M A (2010) Automating the process of taxonomy creation and comparison of taxonomy structures, in: 19th Annual Research Workshop on Strategic and Emerging Technologies, American Accounting Association. San Francisco, California, USA

Cong Y, Kogan A, Vasarhelyi M A (2007) Extraction of Structure and Content from the Edgar Database: A Template-Based Approach, in: Journal of Emerging Technologies in Accounting (JETA) 4(1)/2007, 69-86

Davis A K, Tama-Sweet I (2012) Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases versus MD&A, in: Contemporary Accounting Research (CAR), 29(3)/2012, 804-837

Ditter D, Henselmann K, Scherr E (2011) Using XBRL Technology to Extract Competitive Information from Financial Statements, in: Journal of Intelligence Studies in Business (JISIB), 1/2011, 19-28

Engelberg J, Sankaraguruswamy S (2007) How to Gather Data Using a Web Crawler: An Application Using SAS to Search Edgar. Working Paper, SSRN

Filer Manual (2016) Filer Manual – Volume II EDGAR Filing, Available online on URL: https://www.sec.gov/info/edgar/edgarfm-vol2-v37.pdf

Friedl J E F (2006) Mastering Regular Expressions, Third Edition, O´Reilly Media, Inc., Sebastopol, California, USA

Gaizauskas R, Humphreys K, Azzam S, Wilks Y (1997) Concepticons vs. Lexicons: an Architecture for Multilingual Information Extraction, in: M T Pazienza, Ed. Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology. Springer-Verlag, Berlin Heidelberg, Germany

Garcia D, Norli O (2012) Crawling EDGAR, in: The Spanish Review of Financial Economics (SRFE), 10/2012, 1-10

Gerdes J Jr (2003) EDGAR-Analyzer: automating the analysis of corporate data contained in the SECs EDGAR database, in: Decision Support Systems 35/2003, 7-29

Goyvaerts J (2007) Regular Expressions: The Complete Tutorial, Available online on URL: https://www.princeton.edu/~mlovett/reference/Regular-Expressions.pdf

Goyvaerts J, Levithan S (2012) Regular Expressions Cookbook, Second Edition, O´Reilly Media, Inc., Sebastopol, California, USA

Grant G H, Conlon S J (2006) EDGAR Extraction System: An Automated Approach to Analyze Employee Stock Option Disclosures, in: Journal of Information Systems (JIS), 20(2)/2006, 119-142

Hernandez M A, Ho H, Koutrika G, Krishnamurthy R, Popa L, Stanoi I R, Vaithyanathan S, Das S (2010) Unleashing the Power of Public Data for Financial Risk Measurement, Regulation, and Governance. IBM Technical Report #RJ10475

Jegadeesh N, Wu D (2013) Word power: A new approach for content analysis, in: Journal of Financial Economics (JFE), 110(3)/2013, 712-729

Kambil A, Ginsburg M (1998) Public Access Web Information Systems: Lessons from the Internet EDGAR Project, in: Communications of the ACM (CACM), 41(7)/1998, 91-97

Loughran T, McDonald B (2011a) When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks, in: The Journal of Finance (JoF), 66(1)/2011, 35-65

Loughran T, McDonald B (2011b) Internet Appendix for "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks", Available online on URL: http://www.afajof.org/SpringboardWebApp/userfiles/afa/file/Supplements%20and%20Data%20Sets/Internet%20Appendix%20for%20When%20Is%20a%20Liability%20Not%20a%20Liability%20Textual%20Analysis,%20Dictionaries,%20and%2010-Ks%206989-IA-Feb-2011.pdf

Loughran T, McDonald B (2014) Measuring Readability in Financial Disclosures, in: The Journal of Finance (JoF), 69(4)/2014, 1643-1671

Loughran T, McDonald B (2016) Textual Analysis in Accounting and Finance: A Survey, in: Journal of Accounting Research (JAR), 54(4)/2016, 1187-1230

Mooney R J, Bunescu R (2005) Mining Knowledge from Text Using Information Extraction, in: SIGKDD Explorations (SIGKDD), 7(1)/2005, 3-10

O´Riain S (2012) Semantic Paths in Business Filings Analysis. Ph.D. thesis, National University of Ireland, Galway, Ireland

Pagell R A (1995) EDGAR: Electronic Data Gathering and Receiving, in: Business Information Review (BIR), 11(3)/1995, 56-68

Palmer D D (2010) Text Preprocessing, in: N Indurkhya, F J Damerau, Eds. Handbook of Natural Language Processing, Second Edition, Taylor & Francis Group, Boca Raton, Florida, USA

SEC (1934) Securities Exchange Act of 1934, Available online on URL: https://www.sec.gov/about/laws/sea34.pdf

SEC (2000) HTML Specifications for EDGAR Rel. 7.0, Available online on URL: https://www.sec.gov/info/edgar/ednews/edhtml.htm

SEC (2006) Electronic Filing and the EDGAR System: A Regulatory Overview, Available online on URL: https://www.sec.gov/info/edgar/regoverview.htm

SEC (2010) Important Information about EDGAR, Available online on URL: https://www.sec.gov/edgar/aboutedgar.htm

SEC (2011) Fast Answers – How to Read a 10-K, Available online on URL: https://www.sec.gov/answers/reada10k.htm

SEC (2013) What We Do, Available online on URL: https://www.sec.gov/about/whatwedo.shtml

SEC (2015) Information for FTP Users, Available online on URL: https://www.sec.gov/edgar/searchedgar/ftpusers.htm

SEC EDGAR (2016) EDGAR Public Dissemination Service (PDS) System, Available online on URL: https://www.sec.gov/oit/announcement/public-dissemination-service-system-contact.html

SEC EDGAR (2015) Public Dissemination Service (PDS) Technical Specification, Available online on URL: https://www.sec.gov/info/edgar/specifications/pds_dissemination_spec.pdf

SEC EDGAR Archives (2016) Coca Cola Company´s Financial Statement Submissions on 2016-02-25, Available online on URL: https://www.sec.gov/Archives/edgar/data/21344/000002134416000050/0000021344-16-000050-index.htm

SEC Form Glossary (2015) Index to Forms, Available online on URL: https://www.sec.gov/info/edgar/forms/edgform.pdf

SEC Index Files (2016) Full Index Files, Available online on URL: ftp://ftp.sec.gov/edgar/full-index/

SEC Regulation S-K (2016) Standard Instructions for filing Forms under Securities Act of 1933, Securities Exchange Act of 1934 and Energy Policy and Conservation Act of 1975-Regulation S-K, Available online on URL: http://www.ecfr.gov/cgi-bin/text-idx?SID=8e0ed509ccc65e983f9eca72ceb26753&node=17:3.0.1.1.11&rgn=div5#se17.3.229_1101

SEC Regulation S-T (2016) General Rules and Regulations for electronic Filings, Available online on URL: http://www.ecfr.gov/cgi-bin/text-idx?node=17:3.0.1.1.14&rgn=div5#se17.3.232_1100

SEC Regulation S-X (2016) Form and Content of and Requirements for Financial Statements; Securities Act of 1933, Securities Exchange Act of 1934, Investment Company Act of 1940, Investments Advisers Act of 1940, and Energy Policy and Conservation Act of 1975- Regulation S-X, Available online on URL: http://www.ecfr.gov/cgi-bin/text-idx?SID=8e0ed509ccc65e983f9eca72ceb26753&node=17:3.0.1.1.8&rgn=div5#se17.3.210_11_601

SEC Release 33-8099 (2002) Mandated EDGAR Filing for Foreign Issuers, Available online on URL: https://www.sec.gov/rules/final/33-8099.htm

SEC Release 34-36997 (1996) EDGAR Phase-in Complete on May 6, 1996, Available online on URL: https://www.sec.gov/info/edgar/ednews/34-36997.htm

Srivastava R P (2016) Textual Analysis and Business Intelligence in Big Data Environment: Search Engine versus XBRL, in: Indian Accounting Review (IAR), 20(1)/2016, 1-20

Stümpert T (2008) Extracting Financial Data from SEC Filings for US GAAP Accountants, in: D Seese, C Weinhardt, F Schlottmann, Eds. Handbook on Information Technology in Finance. Springer-Verlag, Berlin Heidelberg, Germany

Stümpert T, Seese D, Centinkaya Ö, Spöth R (2004) EASE – a software agent that extracts financial data from the SEC´s EDGAR database, in: Proceedings of the 4th International ICSC Symposium on Engineering of Intelligent Systems (EIS 2004). Funchal, Portugal

Tetlock P C (2007) Giving Content to Investor Sentiment: The Role of the Media in the Stock Market, in: The Journal of Finance (Jof), 62(3)/2007, 1139-1168

Thai V, Davis B, O´Riain S, O´Sullivan D, Handschuh S (2008) Semantically Enhanced Passage Retrieval for Business Analysis Activity, in: Proceedings of the 16th European Conference on Information Systems (ECIS 2008). Galway, Ireland

Thompson K (1968) Regular Expression Search Algorithm, in: Communications of the ACM (CACM), 11(6)/1968, 419-422

Wilks Y (1997) Information Extraction as a Core Language Technology, in: M T Pazienza, Ed. Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology. Springer-Verlag, Berlin Heidelberg, Germany

W3 Schools (2016), HTML Character Sets, Available online on URL: http://www.w3schools.com/charsets/default.asp

W3C Recommendation (1999) HTML 4.01 Specification, Available online on URL: https://www.w3.org/TR/html401/cover.html

W3C Strict DTD (1999) HTML 4.01 Strict DTD, Available online on URL: https://www.w3.org/TR/html4/strict.dtd

Appendix A: Form Index Files

Year	Q.	Uniform Resource Locator (URL)
1993	4	ftp://ftp.sec.gov/edgar/full-index/1993/QTR4/form.idx
1993	3	ftp://ftp.sec.gov/edgar/full-index/1993/QTR3/form.idx
1993	2	ftp://ftp.sec.gov/edgar/full-index/1993/QTR2/form.idx
1993	1	ftp://ftp.sec.gov/edgar/full-index/1993/QTR1/form.idx
1994	4	ftp://ftp.sec.gov/edgar/full-index/1994/QTR4/form.idx
1994	3	ftp://ftp.sec.gov/edgar/full-index/1994/QTR3/form.idx
1994	2	ftp://ftp.sec.gov/edgar/full-index/1994/QTR2/form.idx
1994	1	ftp://ftp.sec.gov/edgar/full-index/1994/QTR1/form.idx
1995	4	ftp://ftp.sec.gov/edgar/full-index/1995/QTR4/form.idx
1995	3	ftp://ftp.sec.gov/edgar/full-index/1995/QTR3/form.idx
1995	2	ftp://ftp.sec.gov/edgar/full-index/1995/QTR2/form.idx
1995	1	ftp://ftp.sec.gov/edgar/full-index/1995/QTR1/form.idx
1996	4	ftp://ftp.sec.gov/edgar/full-index/1996/QTR4/form.idx
1996	3	ftp://ftp.sec.gov/edgar/full-index/1996/QTR3/form.idx
1996	2	ftp://ftp.sec.gov/edgar/full-index/1996/QTR2/form.idx
1996	1	ftp://ftp.sec.gov/edgar/full-index/1996/QTR1/form.idx
1997	4	ftp://ftp.sec.gov/edgar/full-index/1997/QTR4/form.idx
1997	3	ftp://ftp.sec.gov/edgar/full-index/1997/QTR3/form.idx
1997	2	ftp://ftp.sec.gov/edgar/full-index/1997/QTR2/form.idx
1997	1	ftp://ftp.sec.gov/edgar/full-index/1997/QTR1/form.idx
1998	4	ftp://ftp.sec.gov/edgar/full-index/1998/QTR4/form.idx
1998	3	ftp://ftp.sec.gov/edgar/full-index/1998/QTR3/form.idx
1998	2	ftp://ftp.sec.gov/edgar/full-index/1998/QTR2/form.idx
1998	1	ftp://ftp.sec.gov/edgar/full-index/1998/QTR1/form.idx
1999	4	ftp://ftp.sec.gov/edgar/full-index/1999/QTR4/form.idx
1999	3	ftp://ftp.sec.gov/edgar/full-index/1999/QTR3/form.idx
1999	2	ftp://ftp.sec.gov/edgar/full-index/1999/QTR2/form.idx
1999	1	ftp://ftp.sec.gov/edgar/full-index/1999/QTR1/form.idx
2000	4	ftp://ftp.sec.gov/edgar/full-index/2000/QTR4/form.idx
2000	3	ftp://ftp.sec.gov/edgar/full-index/2000/QTR3/form.idx
2000	2	ftp://ftp.sec.gov/edgar/full-index/2000/QTR2/form.idx
2000	1	ftp://ftp.sec.gov/edgar/full-index/2000/QTR1/form.idx
2001	4	ftp://ftp.sec.gov/edgar/full-index/2001/QTR4/form.idx
2001	3	ftp://ftp.sec.gov/edgar/full-index/2001/QTR3/form.idx
2001	2	ftp://ftp.sec.gov/edgar/full-index/2001/QTR2/form.idx
2001	1	ftp://ftp.sec.gov/edgar/full-index/2001/QTR1/form.idx
2002	4	ftp://ftp.sec.gov/edgar/full-index/2002/QTR4/form.idx
2002	3	ftp://ftp.sec.gov/edgar/full-index/2002/QTR3/form.idx
2002	2	ftp://ftp.sec.gov/edgar/full-index/2002/QTR2/form.idx
2002	1	ftp://ftp.sec.gov/edgar/full-index/2002/QTR1/form.idx
2003	4	ftp://ftp.sec.gov/edgar/full-index/2003/QTR4/form.idx
2003	3	ftp://ftp.sec.gov/edgar/full-index/2003/QTR3/form.idx
2003	2	ftp://ftp.sec.gov/edgar/full-index/2003/QTR2/form.idx
2003	1	ftp://ftp.sec.gov/edgar/full-index/2003/QTR1/form.idx
2004	4	ftp://ftp.sec.gov/edgar/full-index/2004/QTR4/form.idx
2004	3	ftp://ftp.sec.gov/edgar/full-index/2004/QTR3/form.idx
2004	2	ftp://ftp.sec.gov/edgar/full-index/2004/QTR2/form.idx
2004	1	ftp://ftp.sec.gov/edgar/full-index/2004/QTR1/form.idx
2005	4	ftp://ftp.sec.gov/edgar/full-index/2005/QTR4/form.idx
2005	3	ftp://ftp.sec.gov/edgar/full-index/2005/QTR3/form.idx
2005	2	ftp://ftp.sec.gov/edgar/full-index/2005/QTR2/form.idx
2005	1	ftp://ftp.sec.gov/edgar/full-index/2005/QTR1/form.idx
2006	4	ftp://ftp.sec.gov/edgar/full-index/2006/QTR4/form.idx
2006	3	ftp://ftp.sec.gov/edgar/full-index/2006/QTR3/form.idx
2006	2	ftp://ftp.sec.gov/edgar/full-index/2006/QTR2/form.idx
2006	1	ftp://ftp.sec.gov/edgar/full-index/2006/QTR1/form.idx
2007	4	ftp://ftp.sec.gov/edgar/full-index/2007/QTR4/form.idx
2007	3	ftp://ftp.sec.gov/edgar/full-index/2007/QTR3/form.idx
2007	2	ftp://ftp.sec.gov/edgar/full-index/2007/QTR2/form.idx
2007	1	ftp://ftp.sec.gov/edgar/full-index/2007/QTR1/form.idx
2008	4	ftp://ftp.sec.gov/edgar/full-index/2008/QTR4/form.idx
2008	3	ftp://ftp.sec.gov/edgar/full-index/2008/QTR3/form.idx
2008	2	ftp://ftp.sec.gov/edgar/full-index/2008/QTR2/form.idx
2008	1	ftp://ftp.sec.gov/edgar/full-index/2008/QTR1/form.idx
2009	4	ftp://ftp.sec.gov/edgar/full-index/2009/QTR4/form.idx
2009	3	ftp://ftp.sec.gov/edgar/full-index/2009/QTR3/form.idx
2009	2	ftp://ftp.sec.gov/edgar/full-index/2009/QTR2/form.idx
2009	1	ftp://ftp.sec.gov/edgar/full-index/2009/QTR1/form.idx
2010	4	ftp://ftp.sec.gov/edgar/full-index/2010/QTR4/form.idx
2010	3	ftp://ftp.sec.gov/edgar/full-index/2010/QTR3/form.idx
2010	2	ftp://ftp.sec.gov/edgar/full-index/2010/QTR2/form.idx
2010	1	ftp://ftp.sec.gov/edgar/full-index/2010/QTR1/form.idx
2011	4	ftp://ftp.sec.gov/edgar/full-index/2011/QTR4/form.idx
2011	3	ftp://ftp.sec.gov/edgar/full-index/2011/QTR3/form.idx
2011	2	ftp://ftp.sec.gov/edgar/full-index/2011/QTR2/form.idx
2011	1	ftp://ftp.sec.gov/edgar/full-index/2011/QTR1/form.idx
2012	4	ftp://ftp.sec.gov/edgar/full-index/2012/QTR4/form.idx
2012	3	ftp://ftp.sec.gov/edgar/full-index/2012/QTR3/form.idx
2012	2	ftp://ftp.sec.gov/edgar/full-index/2012/QTR2/form.idx
2012	1	ftp://ftp.sec.gov/edgar/full-index/2012/QTR1/form.idx
2013	4	ftp://ftp.sec.gov/edgar/full-index/2013/QTR4/form.idx
2013	3	ftp://ftp.sec.gov/edgar/full-index/2013/QTR3/form.idx
2013	2	ftp://ftp.sec.gov/edgar/full-index/2013/QTR2/form.idx
2013	1	ftp://ftp.sec.gov/edgar/full-index/2013/QTR1/form.idx
2014	4	ftp://ftp.sec.gov/edgar/full-index/2014/QTR4/form.idx
2014	3	ftp://ftp.sec.gov/edgar/full-index/2014/QTR3/form.idx
2014	2	ftp://ftp.sec.gov/edgar/full-index/2014/QTR2/form.idx
2014	1	ftp://ftp.sec.gov/edgar/full-index/2014/QTR1/form.idx
2015	4	ftp://ftp.sec.gov/edgar/full-index/2015/QTR4/form.idx
2015	3	ftp://ftp.sec.gov/edgar/full-index/2015/QTR3/form.idx
2015	2	ftp://ftp.sec.gov/edgar/full-index/2015/QTR2/form.idx
2015	1	ftp://ftp.sec.gov/edgar/full-index/2015/QTR1/form.idx
2016	2	ftp://ftp.sec.gov/edgar/full-index/2016/QTR2/form.idx
2016	1	ftp://ftp.sec.gov/edgar/full-index/2016/QTR1/form.idx

Appendix B: Form Index Files - Descriptive Statistics

Sub. Type	Total	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
1	**33**	0	0	0	0	0	0	0	0	0	6	1	4	2	3	5	1	4	0	2	2	0	2	1	0
3	**538228**	0	0	0	1149	2983	4317	5282	6852	5327	8694	27814	47801	47356	46869	52898	37477	31171	33279	32264	30265	31834	36150	32839	15607
4	**5850937**	0	0	0	4312	12758	18432	19899	22606	21689	56942	324449	476172	474215	468480	486241	456300	387113	404831	393744	398395	400134	400454	397908	225863
5	**186884**	0	0	0	301	2189	3446	4741	5346	5632	6560	16067	19881	16593	15040	13070	11495	10620	9374	9003	8427	8084	8097	6991	5927
25	**3642**	0	0	0	0	0	0	1	0	16	406	468	445	425	335	362	325	205	118	111	108	102	89	76	50
26	**1**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
144	**12111**	0	0	0	9	32	127	107	128	99	193	775	1060	930	1082	1200	687	672	773	790	829	988	856	530	244
425	**81918**	0	0	0	0	0	0	0	4303	4849	4271	4176	4878	6601	4971	5834	3736	4525	4339	4932	3020	4295	6732	7304	3152
487	**15478**	0	346	352	393	394	461	537	552	512	460	473	441	498	564	679	723	874	1042	1074	1101	1063	1130	1187	622
497	**365987**	0	6285	10675	13227	13738	14624	13905	14883	19491	15718	16738	20578	18592	19068	17563	19155	18238	18325	15496	16586	19076	16923	17153	9950
1/A	**413**	0	0	0	0	0	0	0	0	0	1	0	0	2	1	0	0	2	0	0	31	94	135	136	11
10-12B	**578**	0	4	4	37	37	25	29	29	23	19	31	12	15	19	33	39	21	19	27	22	31	38	45	19
10-12B/A	**1318**	0	1	2	69	50	44	50	69	48	40	30	30	15	44	69	106	38	64	67	66	70	150	138	58
10-12G	**2106**	0	2	3	49	76	94	148	84	58	57	38	43	58	75	63	271	159	225	143	118	103	99	92	48
10-12G/A	**3861**	0	4	6	91	146	136	234	235	114	107	80	63	104	116	91	337	274	423	414	262	222	201	134	67
10-C	**1303**	0	82	396	745	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10-C/A	**25**	0	2	12	9	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10-D	**45492**	0	0	0	0	0	0	0	0	0	0	0	0	3	6950	8081	2412	891	934	1775	2983	4571	6035	7097	3760
10-D/A	**2040**	0	0	0	0	0	0	0	0	0	0	0	0	0	483	720	370	54	23	20	55	124	84	69	38
10-K	**167599**	4	1912	2218	4315	6698	6930	6761	6652	6248	6759	8468	8567	9017	8852	8574	8746	9839	9165	8840	8393	8105	8084	7985	6467
10-K/A	**38888**	2	616	933	1495	2152	1943	1798	1530	1578	2010	2021	2096	2180	1510	1470	1801	2320	2213	1995	1840	1765	1557	1258	805
10-K405	**21277**	0	11	1018	1944	3201	3357	3361	3217	3000	2168	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10-K405/A	**3349**	0	5	190	317	538	589	610	483	494	123	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10KSB	**36912**	0	1	95	1052	2013	2001	2173	2856	3199	3419	3410	3557	3458	3399	3513	2686	80	0	0	0	0	0	0	0
10-KSB	**1**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10KSB/A	**11909**	0	0	63	494	713	582	518	724	833	858	1024	955	1380	1382	1009	1284	90	0	0	0	0	0	0	0
10-KSB/A	**2**	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10KSB40	**3441**	0	0	24	213	540	532	547	661	617	307	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10KSB40/A	**625**	0	0	7	42	103	129	110	98	116	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10-KT	**418**	0	1	6	8	18	14	9	12	18	22	25	38	8	19	13	11	11	21	27	23	22	41	22	29
10-KT/A	**106**	0	0	2	1	5	1	3	1	6	6	6	0	3	5	4	3	2	9	9	5	17	8	5	5
10KT405	**86**	0	0	3	14	8	7	10	12	20	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10KT405/A	**17**	0	0	3	6	3	1	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10-Q	**522906**	7	6628	14120	25749	28996	29239	28687	28282	25997	24084	21866	20872	20670	20038	20047	27040	27817	26550	25676	24186	23250	22876	22166	8063
10-Q/A	**41651**	0	558	1692	2319	2036	2135	2426	1926	1722	1778	1694	1704	1946	1574	1065	1517	2195	2060	3539	3045	1829	1420	1095	376
10QSB	**120120**	0	15	1297	5670	7448	7455	8419	11619	11581	11160	10348	10671	10209	10191	10703	3334	0	0	0	0	0	0	0	0
10-QSB	**4**	0	0	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10QSB/A	**17117**	0	1	198	884	984	805	883	1325	1428	1464	1527	1355	1840	2051	1585	787	0	0	0	0	0	0	0	0
10-QSB/A	**1**	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10-QT	**197**	0	3	9	8	18	12	13	17	12	25	7	5	6	10	7	6	5	4	0	5	8	7	7	3
10-QT/A	**33**	0	1	2	1	10	3	1	2	0	0	3	1	0	1	0	1	0	2	1	1	2	0	1	0
10SB12B	**119**	0	0	0	9	12	5	22	12	11	7	3	5	12	13	6	2	0	0	0	0	0	0	0	0
10SB12B/A	**124**	0	0	0	6	8	8	31	31	25	3	3	0	1	4	4	0	0	0	0	0	0	0	0	0
10SB12G	**4105**	0	0	0	71	113	175	1048	1140	348	230	109	138	170	233	274	56	0	0	0	0	0	0	0	0
10SB12G/A	**5523**	0	0	0	79	167	167	1037	1672	694	480	282	184	250	235	253	23	0	0	0	0	0	0	0	0
11-K	**36426**	0	811	611	809	949	1127	1264	1411	1550	2412	2476	2369	2251	2162	2058	1913	1824	1711	1626	1569	1508	1436	1379	1200
11-K/A	**911**	0	37	22	21	26	34	38	33	25	47	95	129	45	43	44	42	45	30	30	25	16	13	58	13
11-KT	**88**	1	4	1	6	14	5	5	1	5	5	10	2	3	8	3	5	0	2	1	2	2	2	1	0
11-KT/A	**2**	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
12G3-2B	**609**	0	0	0	0	0	0	0	3	9	66	66	86	92	109	100	78	0	0	0	0	0	0	0	0
12G32BR	**178**	0	0	0	0	0	0	0	0	7	32	45	26	25	24	15	4	0	0	0	0	0	0	0	0
13FCONP	**1249**	6	5	4	10	13	18	33	92	145	143	133	103	85	41	70	99	72	43	112	20	1	0	0	1
13FCONP/A	**90**	0	4	10	12	15	20	15	5	2	0	6	1	0	0	0	0	0	0	0	0	0	0	0	0
13F-E	**1188**	0	193	232	222	212	252	77	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13F-E/A	**120**	0	5	28	7	15	25	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13F-HR	**193463**	4	4	8	5	8	34	5079	7428	8180	8212	8416	8821	9550	10418	11301	12299	11996	11687	12630	13246	13829	15239	16562	8507
13F-HR/A	**20575**	0	0	0	3	1	0	474	770	979	1074	1055	1228	1199	1353	1435	1803	1109	1234	1525	1264	1300	949	1278	542
13F-NT	**64392**	0	0	0	0	0	0	968	1657	1811	1941	2194	2397	2795	3215	3678	4223	4371	4822	5026	5098	5535	5586	6049	3026
13F-NT/A	**1514**	0	0	0	0	0	0	27	37	44	100	134	182	100	125	112	115	122	57	74	109	84	16	45	31
144/A	**455**	0	0	0	0	6	50	14	24	11	8	25	37	42	38	49	13	14	34	18	12	19	24	12	5
15-12B	**4302**	0	8	23	49	122	201	208	207	213	179	188	209	205	262	312	178	209	224	214	207	253	227	290	114
15-12B/A	**88**	0	0	0	0	3	7	4	10	1	3	4	5	3	6	3	6	5	5	2	5	3	4	7	2
15-12G	**12059**	0	64	158	451	568	707	772	732	730	674	718	731	755	655	683	590	589	425	433	469	382	310	316	147
15-12G/A	**343**	0	1	4	16	14	15	35	24	6	16	17	17	23	23	11	18	25	7	16	19	15	11	5	5
15-15D	**17872**	0	32	104	144	427	715	739	618	575	891	1073	1106	2109	2233	2025	1397	499	270	353	963	471	231	640	257
15-15D/A	**302**	0	0	1	3	5	9	11	9	0	1	48	8	29	13	5	7	9	3	12	5	10	5	109	0
15F-12B	**160**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	63	22	14	11	10	5	4	13	8	10
15F-12B/A	**11**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	1	0	0	1	0	1	1	0	0
15F-12G	**208**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	51	47	28	16	14	15	18	10	7	2
15F-12G/A	**10**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	4	1	0	2	1	0	0	0
15F-15D	**133**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28	14	9	18	15	15	3	15	13	3
15F-15D/A	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
18-12B	**5**	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	2	0	0
18-K	**562**	0	0	0	0	0	0	1	0	1	41	40	35	37	41	36	40	41	44	40	39	39	35	37	15
18-K/A	**1710**	0	0	0	0	0	0	0	1	2	92	110	98	101	111	104	98	132	136	133	137	127	122	130	76
19B-4	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
19B-4E	**20037**	0	0	0	0	0	0	0	0	7	23	18	127	109	271	1481	1679	3127	4113	2184	1555	1520	1403	1830	590
1-A	**494**	0	0	0	0	0	0	0	0	0	23	34	21	30	32	28	26	24	25	20	19	28	55	67	62
1-A POS	**43**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	40
1-A/A	**1194**	0	0	0	0	0	0	0	0	0	100	54	67	45	57	89	60	60	65	54	82	76	105	137	143
1-A-W	**26**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	10	15
1-A-W/A	**5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1
1-E	**137**	0	0	0	0	0	0	0	0	0	5	11	42	41	11	11	8	1	2	0	0	3	1	1	0
1-E AD	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
1-E/A	**70**	0	0	0	0	0	0	0	0	0	0	3	12	15	12	11	12	2	0	1	0	0	0	1	1
1-K	**5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5

1-SA	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
1-U	**31**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	28
1-Z	**8**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	3
1-Z/A	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
20-F	**13313**	0	0	0	9	39	108	144	195	281	1131	1031	1004	1028	980	883	820	768	746	748	723	702	676	691	606		
20-F/A	**3016**	0	0	0	7	3	28	21	55	57	148	189	234	331	233	214	193	188	217	226	267	197	90	81	37		
20FR12B	**124**	0	0	0	2	1	15	11	7	3	17	12	9	5	5	4	4	4	3	3	5	4	4	4	2		
20FR12B/A	**118**	0	0	0	1	0	12	4	7	3	12	4	14	3	0	10	1	0	3	3	11	12	5	4	9		
20FR12G	**358**	0	0	0	4	9	6	24	20	16	41	26	44	35	27	29	13	13	15	6	8	5	11	6	0		
20FR12G/A	**577**	0	0	0	1	10	5	20	39	45	69	49	48	55	33	56	20	17	41	14	19	5	18	13	0		
24F-1	**6**	0	0	0	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
24F-2EL	**20**	0	0	5	4	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
24F-2EL/A	**2**	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
24F-2NT	**149385**	2	3874	4970	6344	6488	7973	8181	8244	8657	7406	6661	5713	5508	5517	5710	5826	6347	6188	6517	7126	6946	7281	7131	4775		
24F-2NT/A	**2992**	0	10	80	79	82	1215	41	60	106	83	118	53	51	79	104	94	119	32	69	40	134	60	96	187		
24F-2TM	**603**	0	52	207	197	147	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
25/A	**33**	0	0	0	0	0	0	0	0	0	0	0	0	1	10	10	6	3	0	1	0	0	1	1	0		
253G1	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1		
253G2	**47**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	43		
25-NSE	**14150**	0	0	0	0	0	0	0	0	0	0	0	0	0	1086	1944	1494	1252	1260	1110	1484	1334	1226	1268	692		
25-NSE/A	**378**	0	0	0	0	0	0	0	0	0	0	0	0	0	52	46	42	38	22	26	20	54	18	54	6		
2-A	**50**	0	0	0	0	0	0	0	0	0	8	8	2	4	3	0	1	1	2	1	3	4	10	3	0		
2-A/A	**80**	0	0	0	0	0	0	0	0	0	7	12	9	7	8	3	8	6	7	6	3	1	2	1	0		
2-AF	**36**	0	0	0	0	0	0	0	0	0	4	7	4	3	0	7	2	1	2	3	1	2	0	0	0		
2-E	**85**	0	0	0	0	0	0	0	0	0	2	2	8	36	18	9	5	4	1	0	0	0	0	0	0		
2-E/A	**11**	0	0	0	0	0	0	0	0	0	0	0	0	3	4	1	2	1	0	0	0	0	0	0	0		
3/A	**33661**	0	0	0	28	162	197	167	317	254	362	1990	3539	3524	3084	3026	2697	2791	2002	1677	1689	1617	1788	1796	954		
305B2	**2689**	0	0	6	16	43	73	59	50	88	90	77	90	104	101	142	77	452	183	136	213	228	249	164	48		
305B2/A	**63**	0	0	0	3	1	0	0	3	0	1	0	13	0	0	0	0	40	0	2	0	0	0	0	0		
34-12H	**6**	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1	0		
35-APP	**6**	0	0	1	1	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
35-APP/A	**2**	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
35-CERT	**5212**	0	422	494	489	480	431	448	425	436	436	402	375	356	18	0	0	0	0	0	0	0	0	0	0		
35-CERT/A	**98**	0	5	10	8	7	8	11	9	15	9	6	0	6	4	0	0	0	0	0	0	0	0	0	0		
39-304D	**9**	0	0	0	0	0	0	0	0	0	1	1	4	1	1	1	0	0	0	0	0	0	0	0	0		
39-304D/A	**1**	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
39-310B	**2**	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
4/A	**197612**	0	0	0	123	236	540	478	705	954	2617	14998	21649	18724	19403	18274	16162	12582	11910	11087	10615	10055	10014	10718	5768		
40-17F1	**277**	0	0	0	0	0	10	8	12	7	10	9	9	11	12	2	7	5	7	5	44	24	25	54	16		
40-17F1/A	**8**	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7		
40-17F2	**17182**	0	0	0	0	75	415	686	806	809	731	1361	1302	1135	1801	1330	1256	737	860	909	800	832	593	597	147		
40-17F2/A	**130**	0	0	0	0	1	5	6	28	3	9	14	9	3	0	20	1	6	1	2	2	1	14	5	0		
40-17G	**42397**	0	0	0	0	0	1	2	1	0	1699	2245	2337	2951	3007	4796	3448	3300	2920	3099	2893	3115	2684	2567	1332		
40-17G/A	**8895**	0	0	0	0	0	0	0	0	0	4	0	11	50	584	1298	520	1012	1723	934	647	652	618	522	320		
40-17GCS	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1		
40-202A	**24**	0	0	0	0	0	0	0	0	0	0	2	0	8	7	0	1	2	0	0	3	0	1	0	0		
40-202A/A	**22**	0	0	0	0	0	0	0	0	0	0	0	2	2	2	1	7	0	2	0	0	0	4	2	0		
40-203A	**2**	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0		
40-203A/A	**2**	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0		
40-205E	**3**	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0		
40-205E/A	**2**	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0		
40-206A	**30**	0	0	0	0	0	0	0	0	0	6	0	0	2	0	0	0	1	8	0	3	4	6	0	0		
40-206A/A	**31**	0	0	0	0	0	0	0	0	0	1	2	0	0	2	2	0	1	1	0	0	1	11	9	1		
40-24B2	**4805**	0	0	0	0	0	0	0	0	0	0	0	0	16	32	1272	803	640	548	316	309	307	240	210	112		
40-24B2/A	**93**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	80	2	7	0	2	1	0	0	0	1		
40-33	**983**	0	0	0	0	0	0	0	0	0	0	127	279	188	54	57	40	41	53	39	12	14	34	18	27		
40-33/A	**10**	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	3	5	0		
40-6B	**64**	0	0	0	0	0	0	0	0	0	4	3	5	7	5	12	13	3	4	2	3	1	0	2	0		
40-6B/A	**585**	0	0	0	0	0	0	0	0	0	7	17	57	9	6	66	84	33	41	57	76	103	26	1	2		
40-6C	**2116**	0	0	0	1	0	0	0	0	0	522	431	178	447	293	244	0	0	0	0	0	0	0	0	0		
40-6C/A	**2766**	0	0	0	0	0	1	0	0	0	877	416	355	279	427	411	0	0	0	0	0	0	0	0	0		
40-8B25	**70**	0	0	0	0	0	0	0	0	0	0	0	1	1	8	6	1	4	8	8	7	9	7	6	4		
40-8F-2	**16**	0	0	0	0	0	0	1	0	2	3	2	5	1	0	0	1	0	0	0	0	0	1	0	0		
40-8F-2/A	**21**	0	0	0	0	0	0	0	1	6	4	1	0	1	1	0	0	2	3	0	0	0	1	1	0		
40-8F-A	**123**	0	0	0	0	0	0	17	22	27	20	15	22	0	0	0	0	0	0	0	0	0	0	0	0		
40-8F-A/A	**65**	0	0	0	0	0	0	2	17	20	7	10	9	0	0	0	0	0	0	0	0	0	0	0	0		
40-8F-B	**4**	0	0	0	0	0	0	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
40-8F-B/A	**1**	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
40-8FC	**3**	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0		
40-8FC/A	**1**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
40-8F-L	**455**	0	0	0	0	0	0	60	71	106	83	83	52	0	0	0	0	0	0	0	0	0	0	0	0		
40-8F-L/A	**180**	0	0	0	0	0	0	18	27	42	38	31	24	0	0	0	0	0	0	0	0	0	0	0	0		
40-8F-M	**536**	0	0	0	0	0	0	30	126	127	103	78	72	0	0	0	0	0	0	0	0	0	0	0	0		
40-8F-M/A	**172**	0	0	0	0	0	0	7	37	45	32	25	26	0	0	0	0	0	0	0	0	0	0	0	0		
40-APP	**6883**	0	0	0	0	0	0	0	2	1	276	304	194	248	215	252	885	637	485	562	440	664	541	826	351		
40-APP/A	**15237**	0	0	0	0	0	0	0	0	0	295	277	266	342	295	278	950	1585	1126	1344	2096	1469	1785	1740	1389		
40-F	**2150**	0	0	0	0	0	0	0	0	1	124	135	162	172	171	165	140	123	127	143	153	145	143	132	114		
40-F/A	**423**	0	0	0	0	0	0	0	0	0	21	18	25	51	46	54	28	42	15	24	36	23	23	14	3		
40FR12B	**95**	0	0	0	0	0	0	0	0	0	7	6	4	7	8	6	5	3	4	15	15	3	3	7	2		
40FR12B/A	**45**	0	0	0	0	0	0	0	0	0	3	1	4	8	2	4	1	0	3	5	6	4	1	3	0		
40FR12G	**22**	0	0	0	0	0	0	0	0	0	1	4	5	2	3	1	0	0	0	3	2	0	1	0	0		
40FR12G/A	**8**	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	2	2	0	0	0	0		
40-OIP	**394**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	61	72	73	41	26	60	16	41	4		
40-OIP/A	**707**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	143	240	58	65	80	58	37	17		
40-RPT	**2**	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0		
424A	**1613**	0	21	46	167	161	100	98	729	44	21	33	46	21	26	30	17	13	12	11	7	4	2	3	1		
424B1	**7984**	0	121	216	904	974	898	775	479	394	687	319	551	239	220	194	99	79	170	183	87	174	91	117	13		
424B2	**137674**	0	1204	1562	1704	2026	1864	1209	1257	1743	1815	2129	1552	1860	5093	7709	7621	5961	8869	12391	13762	13937	15513	16771	10122		
424B3	**254046**	0	3373	4410	4701	6803	7684	7038	6977	8139	7989	10457	14349	14150	16180	15337	11617	12752	17062	19862	17644	14616	15000	12377	5529		
424B4	**10096**	0	180	150	662	757	680	975	658	419	435	349	584	497	444	433	157	146	278	278	292	662	492	386	182		
424B5	**98623**	0	568	869	1452	1983	2649	2185	2126	3144	3521	4846	6197	6922	5690	5453	3145	3845	5263	4079	7748	6349	8383	8204	4002		

424B7	**5978**	0	0	0	10	14	3	0	0	0	0	0	0	23	583	2203	1550	155	192	172	206	267	228	274	98
424B8	**554**	0	0	0	0	0	0	0	0	0	0	0	0	0	45	56	39	10	30	61	70	74	42	54	73
424H	**180**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	170
424H/A	**4**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
485A24E	**84**	0	6	41	23	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
485A24F	**20**	0	0	4	6	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
485APOS	**35276**	0	393	834	1352	1524	2147	3126	1715	1385	1276	1685	1684	2348	1467	1502	1440	1807	1991	1442	1304	1481	1319	1407	647
485B24E	**1726**	0	311	539	584	292	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
485B24F	**60**	0	2	5	15	38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
485BPOS	**151903**	0	4596	5663	7024	6946	6340	6080	7357	6952	6888	6722	5585	5859	5684	5654	5805	5539	5503	8072	8626	8572	8612	8586	5238
485BXT	**14307**	0	0	0	4	36	51	141	128	150	151	233	216	440	363	451	678	776	809	1156	1393	1726	2038	2129	1238
485BXTF	**1**	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
486A24E	**1**	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
486APOS	**121**	0	21	0	0	0	1	8	10	5	7	5	6	10	5	4	5	1	4	3	7	1	12	6	0
486B24E	**2**	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
486BPOS	**484**	0	28	12	6	9	13	13	19	19	30	33	21	12	15	14	17	13	16	24	29	25	34	45	37
497AD	**3146**	0	0	0	0	0	0	0	84	1	9	94	128	250	162	237	290	275	289	210	269	299	295	183	71
497H2	**82**	0	0	1	1	2	7	5	1	8	6	10	2	0	2	4	0	1	0	10	6	16	0	0	0
497J	**102391**	0	1502	2056	2877	3282	3844	3776	4488	4606	4653	4893	4378	4566	4613	4782	5264	4991	4978	5262	5828	5821	6262	6166	3503
497K	**113564**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1110	12051	15508	15656	18426	18646	20106	12061
497K1	**454**	0	0	0	0	0	112	244	22	16	12	13	0	5	9	14	7	0	0	0	0	0	0	0	0
497K2	**186**	0	0	0	0	0	54	92	16	7	5	6	1	1	0	2	2	0	0	0	0	0	0	0	0
497K3A	**63**	0	0	0	0	0	1	21	4	0	0	4	5	4	5	5	13	1	0	0	0	0	0	0	0
497K3B	**2666**	0	0	0	0	0	27	120	314	316	242	240	247	287	295	269	249	60	0	0	0	0	0	0	0
5/A	**7535**	0	0	0	10	32	135	143	141	174	225	660	1308	797	702	586	450	430	314	242	330	274	248	211	123
6B NTC	**34**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	12	6	5	5	3	0	2
6B ORDR	**29**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	11	2	5	5	3	0	1
6-K	**326751**	0	0	3	43	183	442	861	1864	2822	15717	20148	23440	25453	26806	26877	23408	20656	20103	21226	21347	21329	21472	21383	11168
6-K/A	**4274**	0	0	1	2	3	6	11	28	52	122	212	258	316	368	303	320	322	310	333	344	276	266	267	154
8-A12B	**18287**	0	223	299	758	882	1087	806	596	790	985	710	778	641	883	1179	1015	861	773	734	762	1009	913	1221	382
8-A12B/A	**4289**	0	110	127	261	280	291	310	201	196	384	197	165	151	174	207	176	242	147	114	134	150	113	111	48
8A12BEF	**97**	0	0	40	26	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8A12BT	**136**	0	0	105	25	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8A12BT/A	**8**	0	0	1	5	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8-A12G	**9845**	0	29	97	851	1204	999	1038	862	531	483	404	478	471	452	311	271	200	182	251	204	155	158	152	62
8-A12G/A	**2309**	0	28	57	184	214	168	250	181	152	155	141	130	123	95	73	79	64	58	38	38	23	28	24	6
8-B12B	**89**	0	11	22	29	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8-B12B/A	**30**	0	2	15	11	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8-B12G	**58**	0	2	4	15	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8-B12G/A	**13**	0	0	7	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8F-2 NTC	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1
8F-2 ORDR	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
8-K	**1376248**	1	3623	6336	15716	24098	27956	27901	29649	35240	45016	67656	91362	116283	106308	101305	90923	81964	80442	78824	77353	76369	76930	76407	38586
8-K/A	**68142**	0	375	798	2155	2858	3285	2941	2945	2803	3073	3632	4093	5396	4898	4375	3407	2987	2787	3865	2771	2804	2613	2280	1001
8-K12B	**136**	0	0	0	0	0	0	0	0	0	1	0	1	0	7	10	3	11	14	15	10	14	18	20	12
8-K12B/A	**11**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	1	1	0	1	2	0	1	1
8-K12G3	**582**	0	0	0	8	20	44	40	80	49	47	41	40	26	27	27	23	25	17	16	11	16	9	13	3
8-K12G3/A	**113**	0	0	0	0	1	0	3	16	38	18	7	4	4	3	4	0	1	1	3	2	0	4	4	0
8-K15D5	**71**	0	0	1	5	17	2	1	2	4	13	2	3	3	4	5	2	0	0	6	1	0	0	0	0
8-K15D5/A	**2**	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8-M	**47**	0	0	0	0	0	0	0	0	0	3	8	5	4	4	3	2	1	5	4	0	5	0	2	1
9-M	**11**	0	0	0	0	0	0	0	0	0	0	1	2	2	0	2	1	0	1	2	0	0	0	0	0
ABS-15G	**4201**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	650	697	779	1128	947
ABS-15G/A	**1360**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	295	248	292	339	186
ADB	**18**	0	0	0	0	0	0	0	0	0	11	7	0	0	0	0	0	0	0	0	0	0	0	0	0
ADN-MTL	**38**	0	0	0	0	0	0	0	0	1	13	12	0	1	2	0	3	0	4	0	1	1	0	0	0
ADV/A	**1**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ADVCO	**50**	0	0	0	0	0	0	0	0	0	49	1	0	0	0	0	0	0	0	0	0	0	0	0	0
ADV-E	**2417**	0	0	0	0	0	0	0	0	3	433	443	287	237	212	190	190	218	196	4	4	0	0	0	0
ADV-H-C	**5**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	3	0	0	0	0	0	0	0	0
ADV-H-T	**42**	0	0	0	0	0	0	0	0	0	5	5	4	9	2	2	0	0	3	7	2	1	0	0	2
ADV-NR	**333**	0	0	0	0	0	0	0	0	8	6	7	15	58	73	11	4	14	14	19	21	29	24	18	12
ADVW	**8**	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AFDB	**9**	0	0	0	0	0	0	0	0	0	5	4	0	0	0	0	0	0	0	0	0	0	0	0	0
AFDB/A	**1**	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
ANNLRPT	**76**	0	0	0	0	0	0	0	0	0	2	9	3	6	8	3	7	4	6	7	5	5	4	7	0
ANNLRPT/A	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
APP NTC	**905**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	105	122	104	201	105	115	97	56
APP ORDR	**862**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	79	117	93	199	115	101	106	52
APP WD	**882**	0	0	0	0	0	0	0	0	0	35	46	26	23	31	24	97	25	217	52	96	83	29	67	31
APP WD/A	**23**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	6	2	1	0	9	0
APP WDG	**51**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	19	3	11	5
ARS	**36465**	0	40	68	125	147	108	112	81	93	3355	3163	2831	2950	2879	2657	2899	2389	2295	2074	2015	1986	2071	2095	32
ARS/A	**34**	0	0	0	0	0	0	0	0	0	3	2	0	0	0	5	8	3	2	5	0	2	2	1	1
AW	**3355**	0	40	48	58	54	81	64	95	121	100	104	141	213	229	189	224	247	271	345	174	157	132	170	98
AW WD	**88**	0	0	0	0	0	0	0	0	0	3	8	14	5	6	20	3	4	7	9	1	2	1	1	4
BDCO	**423**	0	0	0	0	0	0	0	0	0	91	332	0	0	0	0	0	0	0	0	0	0	0	0	0
BW-2	**27**	0	0	0	2	3	4	4	4	4	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0
BW-3	**93**	0	0	0	4	4	5	14	16	18	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	**47**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	47
C/A	**25**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25
CB	**1962**	0	0	0	0	0	0	0	0	0	105	125	120	416	143	108	72	126	163	130	100	139	91	80	44
CB/A	**1624**	0	0	0	0	0	0	0	0	0	137	115	95	55	104	167	84	137	140	153	129	85	67	122	34
CERT	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
CERTAMX	**1545**	0	0	0	0	0	0	0	0	3	213	148	222	267	231	304	157	0	0	0	0	0	0	0	0
CERTARCA	**341**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	72	113	31	56	31	31	7
CERTBATS	**32**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	5	0	14	11
CERTBSE	**13**	0	0	0	0	0	0	0	0	0	4	4	0	0	4	1	0	0	0	0	0	0	0	0	0
CERTCSE	**2**	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
CERTNAS	**2313**	0	0	0	0	0	0	0	0	0	0	0	0	0	129	310	181	131	213	167	198	259	348	279	98
CERTNYS	**4984**	1	0	0	0	0	0	0	0	7	438	368	363	298	315	319	192	225	362	354	448	490	372	296	136

CERTPAC	**539**	0	0	0	0	0	0	0	0	0	9	0	3	17	12	60	226	148	52	0	10	2	0	0	0
CERTPBS	**3**	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CFPORTAL	**11**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
CFPORTAL/A	**15**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15
CORRESP	**153987**	0	0	0	0	0	0	0	0	0	0	1	916	8923	11777	11581	10573	15088	17927	17242	15887	16028	12858	11609	3577
CT ORDER	**11715**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1095	1494	1592	1429	1488	1329	1343	1287	658
C-W	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
D	**147355**	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	389	14260	17861	18348	18419	20157	22453	23462	12004
D/A	**101613**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	61	7141	11940	12672	13410	14680	15972	16118	9617
DEF 14A	**152216**	0	1925	3823	6104	8639	8566	8517	8511	8041	8052	7632	7401	7644	7542	7156	6917	7222	6143	5796	5554	5614	5526	5635	4256
DEF 14C	**10075**	0	21	49	125	161	205	220	336	421	500	529	694	650	657	658	622	585	647	528	482	559	588	569	269
DEF13E3	**99**	0	1	11	19	27	13	25	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DEF13E3/A	**115**	0	6	11	51	15	24	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DEFA14A	**72777**	0	624	589	1137	1258	1203	1284	1709	1914	1803	2006	2797	2958	5071	5750	4458	5098	4941	4449	4160	5262	5463	5063	3780
DEFA14C	**362**	0	2	5	6	15	6	8	12	14	12	7	11	6	21	44	55	20	17	22	14	16	19	23	7
DEFC14A	**2601**	0	159	47	71	119	95	151	161	119	113	104	78	81	95	117	162	162	109	83	91	125	115	141	103
DEFC14C	**38**	0	6	0	0	0	3	3	3	4	2	4	0	0	0	6	1	0	2	2	0	0	0	2	0
DEFM14A	**4486**	0	29	61	155	175	228	256	301	238	174	184	207	253	285	345	207	137	202	202	149	191	185	219	103
DEFM14C	**375**	0	1	7	12	12	15	38	24	28	26	21	16	12	23	20	16	17	13	19	17	6	13	12	7
DEFN14A	**535**	0	0	5	58	24	22	35	31	16	15	16	15	29	49	53	28	44	23	16	12	12	16	6	10
DEF-OC	**2**	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
DEFR14A	**5403**	0	10	124	214	281	381	333	348	321	371	240	258	279	304	229	249	289	141	187	221	163	146	167	147
DEFR14C	**362**	0	0	0	7	14	9	11	11	13	19	30	32	30	27	30	28	15	14	13	9	23	13	9	5
DEFS14A	**4776**	0	79	196	523	812	746	690	815	712	203	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DEFS14C	**86**	0	0	2	13	23	8	9	13	13	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DEL AM	**994**	0	2	6	12	45	35	32	39	54	42	32	53	50	17	107	55	67	53	39	22	47	112	51	22
DFAN14A	**19408**	0	20	128	600	836	678	480	688	1187	970	835	542	650	1302	1069	1181	1234	1115	657	965	1217	1139	1105	810
DFRN14A	**535**	0	0	4	26	21	68	56	26	68	32	29	9	22	22	24	20	20	18	6	8	18	22	6	10
DOS	**22**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	5
DOS/A	**20**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	5
DOSLTR	**10**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0
DRS	**906**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	61	282	352	201	10
DRS/A	**1167**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24	331	454	292	66
DRSLTR	**745**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	226	318	168	15
DSTRBRPT	**134**	0	0	0	0	0	0	0	0	0	0	12	9	5	6	4	7	4	20	15	8	3	14	15	12
DSTRBRPT/A	**15**	0	0	0	0	0	0	0	0	0	0	8	5	2	0	0	0	0	0	0	0	0	0	0	0
EBRD	**16**	0	0	0	0	0	0	0	0	0	6	10	0	0	0	0	0	0	0	0	0	0	0	0	0
EBRD/A	**1**	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
EFFECT	**95230**	0	0	0	0	0	0	0	0	0	0	0	0	0	5213	10188	6649	7620	10197	11531	10239	11023	9489	9613	3468
F-1	**1025**	0	0	0	6	7	12	25	44	11	39	31	77	66	58	83	38	35	95	65	58	65	113	79	18
F-1/A	**2941**	0	0	0	10	30	26	88	110	28	93	70	149	197	153	251	88	81	360	217	176	202	303	233	76
F-10	**771**	0	0	0	0	0	0	0	0	0	31	29	112	37	78	76	28	48	34	38	49	41	78	64	28
F-10/A	**931**	0	0	0	0	0	0	0	0	0	30	28	142	93	83	83	29	65	27	29	44	52	86	108	32
F-10EF	**5**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	3	0	0	1	0	0	0	0	0	0
F-10POS	**48**	0	0	0	0	0	0	0	0	0	1	4	1	4	8	5	4	2	8	3	1	3	0	1	3
F-1MEF	**105**	0	0	0	0	2	0	1	1	0	4	4	1	7	8	10	0	4	7	8	2	14	14	15	3
F-2	**16**	0	0	0	0	0	1	0	2	0	3	1	7	2	0	0	0	0	0	0	0	0	0	0	0
F-2/A	**37**	0	0	0	2	0	0	1	7	0	8	9	6	4	0	0	0	0	0	0	0	0	0	0	0
F-3	**2567**	0	0	2	3	21	12	14	39	33	105	203	97	79	142	256	81	221	134	87	156	100	214	536	32
F-3/A	**3864**	0	0	2	1	18	15	14	27	17	133	209	181	165	161	261	73	336	136	57	150	81	272	1460	95
F-3ASR	**1441**	0	0	0	0	0	0	0	0	0	0	0	0	5	67	46	149	136	255	59	128	425	52	97	22
F-3D	**39**	0	0	0	0	0	0	0	0	0	3	0	0	0	0	5	4	1	5	3	3	6	1	7	1
F-3DPOS	**21**	0	0	0	0	0	2	1	0	0	1	1	0	4	0	0	1	1	0	2	1	3	0	2	2
F-3MEF	**30**	0	0	0	0	0	0	1	1	7	1	4	1	4	0	0	0	0	4	1	1	3	1	1	0
F-4	**3217**	0	1	0	6	68	40	30	33	66	180	435	240	479	99	178	32	106	177	465	414	75	36	44	13
F-4 POS	**31**	0	0	0	0	0	0	0	0	0	4	7	1	0	4	0	1	2	2	5	2	2	0	0	1
F-4/A	**6312**	0	0	0	7	86	135	33	58	108	320	560	1179	541	101	188	47	118	159	919	1443	134	67	65	44
F-4MEF	**8**	0	0	0	0	0	0	0	0	3	0	0	1	1	0	0	0	0	1	0	0	0	0	1	1
F-6	**2175**	0	0	0	0	0	0	2	2	0	99	67	96	89	77	88	86	141	273	237	162	186	348	164	58
F-6 POS	**1550**	0	0	0	0	0	0	0	0	0	14	37	39	34	47	89	125	125	118	76	152	178	176	222	118
F-6/A	**138**	0	0	0	0	0	0	2	0	0	7	6	11	9	3	12	6	10	24	10	10	10	6	8	4
F-6EF	**8337**	0	0	0	0	0	0	0	0	0	77	46	47	53	96	130	1282	635	580	798	804	982	1359	844	604
F-7	**32**	0	0	0	0	0	0	0	0	0	2	2	0	1	1	2	3	4	1	0	1	2	5	7	1
F-7 POS	**5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	2	0
F-7/A	**8**	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	2	1	1
F-8	**31**	0	0	0	0	0	0	0	0	0	4	3	2	4	4	2	3	3	3	1	0	1	1	0	0
F-8 POS	**8**	0	0	0	0	0	0	0	0	0	2	1	0	0	3	0	0	0	0	0	0	0	2	0	0
F-8/A	**49**	0	0	0	0	0	0	0	0	0	4	3	2	4	10	3	7	1	9	1	1	4	0	0	0
F-80	**24**	0	0	0	0	0	0	0	0	0	4	0	0	3	4	1	0	1	0	3	1	3	1	1	2
F-80/A	**21**	0	0	0	0	0	0	0	0	0	3	0	0	1	4	0	0	2	0	3	1	7	0	0	0
F-80POS	**10**	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	3	0	0	1	2	2
F-9	**125**	0	0	0	0	0	0	0	0	0	10	15	15	6	8	10	14	8	9	14	16	0	0	0	0
F-9 POS	**5**	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	1	0	0	0	0
F-9/A	**110**	0	0	0	0	0	0	0	0	0	10	11	11	5	9	8	11	12	5	13	15	0	0	0	0
F-9EF	**4**	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
F-N	**273**	0	0	0	0	0	0	0	0	0	5	8	10	11	20	15	25	30	15	21	25	16	23	37	12
F-N/A	**19**	0	0	0	0	0	0	0	0	0	2	1	0	0	1	0	0	0	3	1	1	0	5	1	4
FOCUSN	**24550**	0	0	0	0	0	0	0	0	0	268	1743	1633	1706	1706	1760	1636	1530	1691	1748	1876	1905	1945	1980	1423
FOCUSN/A	**833**	0	0	0	0	0	0	0	0	0	27	37	50	36	32	55	47	44	73	82	60	66	31	127	66
FWP	**105002**	0	0	0	0	0	0	0	0	0	0	0	0	477	10147	11954	8304	6491	7614	8930	11484	12094	12827	11003	3677
F-X	**2518**	0	0	0	0	0	0	0	0	1	142	159	172	183	220	230	124	177	194	211	150	147	190	142	76
F-X/A	**167**	0	0	0	0	0	0	0	0	0	13	4	18	17	6	15	16	19	14	10	21	4	3	2	5
G-405	**60**	0	0	1	2	0	1	1	0	0	7	7	5	5	4	5	3	3	3	3	3	2	3	1	1
G-405/A	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0
G-405N	**46**	0	0	0	0	0	1	0	1	0	2	4	4	5	6	3	2	3	3	3	1	2	3	2	1
G-405N/A	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
G-FIN	**18**	0	0	0	0	0	0	0	0	0	1	6	0	0	0	0	2	1	2	2	1	0	1	1	1
G-FIN/A	**222**	0	0	0	0	0	0	0	0	0	18	57	7	8	19	9	14	14	7	17	12	14	16	10	0
G-FINW	**14**	0	0	0	0	0	0	0	0	0	1	1	1	0	0	2	0	0	2	2	1	0	2	0	2
IADB	**13**	0	0	0	0	0	0	0	0	0	9	4	0	0	0	0	0	0	0	0	0	0	0	0	0
ID-NEWCIK	**1**	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0

IFC	**6**	0	0	0	0	0	0	0	0	0	0	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IRANNOTICE	**1788**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	520	537	468	263
MA	**732**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	508	194	30
MA/A	**413**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	26	250	137
MA-A	**544**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	319	224
MA-I	**5539**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3962	989	588
MA-I/A	**2265**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	97	1289	879
MA-W	**63**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	40	20
MSD	**39**	0	0	0	0	0	0	0	0	0	2	4	2	0	3	5	1	2	2	8	3	3	3	3	1	0
MSD/A	**351**	0	0	0	0	0	0	0	0	0	23	25	20	19	25	29	49	35	22	22	18	17	21	21	5
MSDCO	**22**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	0	0	0	0	0	0
MSDW	**77**	0	0	0	0	0	0	0	0	0	4	1	4	3	6	4	20	13	4	6	0	3	3	5	1
N-1	**2**	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
N-1/A	**142**	0	0	12	32	20	15	14	9	9	3	7	9	12	0	0	0	0	0	0	0	0	0	0	0	0
N-14	**2789**	0	2	4	7	21	98	102	172	186	121	172	153	113	131	143	173	266	178	129	139	123	124	175	57
N-14 8C	**197**	0	0	0	0	0	0	13	14	6	8	1	6	13	4	7	13	13	7	21	18	14	20	12	7
N-14 8C/A	**226**	0	0	0	0	0	0	9	14	6	6	3	4	13	5	17	8	22	15	13	30	13	25	15	8
N-14/A	**1875**	0	2	16	15	20	60	60	132	139	89	111	120	111	164	146	101	161	72	59	58	72	68	53	46
N-14AE	**889**	0	0	0	0	18	56	40	131	63	150	101	67	84	136	37	6	0	0	0	0	0	0	0	0
N-14AE/A	**182**	0	0	0	0	1	6	15	12	23	38	28	11	16	23	7	2	0	0	0	0	0	0	0	0
N14AE24	**226**	0	6	80	84	56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N14AE24/A	**57**	0	0	25	21	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N14EL24	**151**	0	9	59	46	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N14EL24/A	**89**	0	18	32	23	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-14MEF	**3**	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
N-18F1	**1208**	0	21	105	107	88	105	84	82	44	37	33	44	58	27	31	63	59	34	31	23	25	81	21	5
N-18F1/A	**30**	0	0	1	1	5	3	0	4	3	1	1	0	1	2	0	2	4	1	1	0	0	0	0	0
N-1A	**1643**	0	12	31	31	69	174	163	149	106	97	81	86	60	88	64	46	40	49	56	62	69	45	43	22
N-1A EL	**352**	0	35	54	144	119	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-1A EL/A	**463**	0	54	88	169	152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-1A/A	**2496**	0	8	38	79	99	241	246	230	113	107	105	92	94	130	117	84	63	80	88	110	110	126	93	43
N-2	**3099**	0	16	32	54	77	122	172	96	142	242	220	222	177	128	166	120	126	176	178	141	163	140	130	59
N-2/A	**4760**	0	32	56	85	99	144	219	116	228	352	366	321	326	179	266	142	138	243	257	301	335	238	243	74
N-23C-1	**920**	0	15	28	41	32	86	114	150	139	132	119	62	2	0	0	0	0	0	0	0	0	0	0	0
N-23C-1/A	**16**	0	0	0	0	0	8	1	0	0	2	3	2	0	0	0	0	0	0	0	0	0	0	0	0
N-23C-2	**1342**	0	0	0	0	2	7	4	6	6	17	18	5	4	17	12	423	323	132	112	54	72	52	30	46
N-23C-2/A	**51**	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	16	25	2	1	0	2	0	1	1
N-23C3A	**1282**	0	0	0	0	0	7	21	51	59	71	75	81	81	81	92	84	73	72	67	61	75	78	103	50
N-23C3A/A	**18**	0	0	0	0	0	0	2	0	2	3	2	0	0	0	1	2	1	2	0	0	1	1	1	0
N-23C3B	**385**	0	0	0	0	12	22	30	35	39	42	34	28	25	24	24	11	9	8	9	9	8	8	6	2
N-23C3B/A	**3**	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
N-23C3C	**45**	0	0	0	0	0	11	18	5	7	1	0	0	0	0	1	0	0	0	0	0	0	1	1	0
N-23C3C/A	**45**	0	0	0	0	0	0	0	4	2	5	5	3	3	8	4	4	5	1	0	0	0	0	0	1
N-27D-1	**25**	0	0	0	0	0	3	3	0	1	2	4	4	4	4	0	0	0	0	0	0	0	0	0	0
N-2MEF	**214**	0	0	0	0	2	2	1	0	10	16	30	27	19	15	26	1	8	7	11	13	14	6	5	1
N-3	**39**	0	0	1	0	1	7	2	4	3	3	1	11	2	0	4	0	0	0	0	0	0	0	0	0
N-3 EL	**5**	0	0	1	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-3 EL/A	**6**	0	0	1	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-3/A	**18**	0	0	0	3	0	4	2	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0
N-30B-2	**9732**	0	328	679	745	534	531	447	316	351	349	295	300	278	370	442	458	489	475	441	421	398	415	430	240
N-30D	**74694**	0	1240	4599	7522	7966	8401	8614	8679	9089	8793	4073	608	396	495	510	495	458	430	434	418	403	438	437	196
N-30D/A	**1114**	0	21	52	96	87	113	138	127	137	187	107	20	3	2	1	0	3	6	4	1	1	0	5	3
N-4	**2057**	0	0	6	6	30	107	149	144	151	168	100	115	56	96	169	104	73	75	124	88	140	105	35	16
N-4 EL	**145**	0	1	30	62	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-4 EL/A	**138**	0	0	24	61	53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-4/A	**1832**	0	1	12	26	25	98	127	157	113	125	122	62	67	77	101	156	77	77	135	82	89	42	43	18
N-5	**10**	0	0	0	3	2	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	2	0	0	0
N-5/A	**9**	0	0	0	1	1	2	0	1	0	0	0	0	0	2	0	0	0	0	0	0	2	0	0	0
N-54A	**244**	0	0	0	6	12	8	3	11	4	6	12	46	21	15	9	7	4	10	13	15	14	12	14	2
N-54A/A	**8**	0	0	0	0	2	1	0	1	0	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0
N-54C	**160**	0	0	0	3	4	7	13	6	9	3	9	6	18	21	12	13	9	10	8	1	0	5	2	1
N-54C/A	**4**	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0
N-6	**583**	0	0	0	0	0	0	0	0	0	28	79	56	22	43	101	94	15	41	7	19	21	44	10	3
N-6/A	**415**	0	0	0	0	0	0	0	0	0	11	50	45	34	38	33	112	30	5	6	15	10	12	12	2
N-6F	**125**	0	0	0	0	0	2	4	1	0	1	4	0	2	5	10	3	3	13	23	8	13	16	14	3
N-6F/A	**60**	0	0	0	0	0	1	2	0	0	0	0	1	1	1	1	0	2	10	22	13	1	2	2	1
N-8A	**3772**	0	51	115	249	266	285	276	223	169	261	207	207	145	156	159	105	112	135	143	130	143	92	102	41
N-8A/A	**544**	0	7	17	19	32	25	62	14	25	14	16	11	22	30	25	13	11	33	7	69	55	23	8	6
N-8B-2	**180**	0	0	13	15	28	30	28	19	13	15	1	0	0	1	1	2	1	0	0	4	2	4	3	0
N-8B-2/A	**165**	0	5	22	6	8	12	6	15	11	12	11	5	0	4	7	14	13	2	4	2	1	3	0	2
N-8B-4	**1**	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-8F	**2416**	0	0	0	0	0	0	0	0	0	0	0	77	309	257	208	227	330	221	132	147	193	141	128	46
N-8F NTC	**1135**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	187	300	138	125	98	98	150	39
N-8F ORDR	**1283**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	192	300	144	127	166	164	153	37
N-8F/A	**1110**	0	0	0	0	0	0	0	0	0	0	0	36	205	131	133	69	167	95	49	73	66	37	24	25
N-CR	**27**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	20
N-CR/A	**6**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
N-CSR	**47853**	0	0	0	0	0	0	0	0	0	0	3178	4408	3834	3668	3637	3667	3542	3285	3336	3242	3326	3386	3449	1895
N-CSR/A	**2514**	0	0	0	0	0	0	0	0	0	0	215	309	342	238	245	248	81	174	129	147	144	79	116	47
N-CSRS	**40102**	0	0	0	0	0	0	0	0	0	0	1083	2912	3004	3067	3106	3116	3093	3137	3126	3172	3209	3304	3342	1431
N-CSRS/A	**799**	0	0	0	0	0	0	0	0	0	0	114	92	62	64	70	79	54	46	54	54	44	28	27	11
N-MFP	**41414**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	699	8551	8033	7586	7321	7032	2192
N-MFP/A	**2240**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	182	907	267	138	351	207	188
N-MFP1	**1039**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1039
N-MFP1/A	**109**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	109
NO ACT	**8514**	0	0	0	1	1	0	0	1	19	797	679	761	600	525	551	743	680	555	455	475	525	483	417	246
N-PX	**37689**	0	0	0	0	0	0	0	0	0	0	1	3581	3607	3429	3409	3442	3000	2940	2870	2841	2856	2829	2870	14
N-PX/A	**1594**	0	0	0	0	0	0	0	0	0	0	0	190	376	112	103	130	100	199	95	70	92	50	67	10
N-Q	**76119**	0	0	0	0	0	0	0	0	0	0	3	2526	6786	6489	6513	6522	6295	6159	6159	6202	6312	6452	6468	3233
N-Q/A	**912**	0	0	0	0	0	0	0	0	0	0	0	47	135	84	103	124	75	77	62	48	46	46	44	21
NRSRO-CE	**12**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	10

NRSRO-CE/A	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
NRSRO-UPD	**36**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	17
NSAR-A	**68163**	0	1832	2802	3220	3276	3232	3239	3344	3369	3301	3296	3288	3212	3085	3126	3061	3011	2911	2886	2837	2815	2834	2830	1356
NSAR-A/A	**3741**	0	113	187	193	169	123	166	233	124	125	100	76	176	105	144	130	116	399	154	146	377	152	149	84
NSAR-AT	**200**	0	24	19	14	12	21	14	16	12	1	7	8	7	12	7	3	5	1	3	2	9	1	1	1
NSAR-AT/A	**16**	0	1	1	0	2	2	0	0	1	0	2	1	0	0	1	0	1	1	0	0	2	0	1	0
NSAR-B	**67801**	1	1773	2619	3148	3170	3226	3303	3299	3459	3350	3314	3272	3184	2971	3207	3075	3020	2899	2818	2807	2820	2816	2790	1460
NSAR-B/A	**5878**	0	171	187	232	245	227	249	309	225	254	214	246	315	198	289	279	169	517	187	259	525	254	239	88
NSAR-BT	**665**	0	51	50	47	47	32	47	34	37	12	37	31	9	23	33	17	9	12	45	17	14	9	39	13
NSAR-BT/A	**54**	0	3	6	0	2	0	4	8	6	0	3	5	3	0	8	1	0	2	1	0	0	0	2	0
NSAR-U	**15145**	0	1	9	62	148	680	740	748	802	817	836	842	835	850	823	823	823	808	773	770	772	739	722	722
NSAR-U/A	**319**	0	0	0	2	30	24	11	24	48	6	21	16	17	20	4	14	12	5	7	13	9	6	17	13
NT 10-D	**7**	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	1	0	0
NT 10-D/A	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
NT 10-K	**39588**	0	51	187	749	1533	1687	1815	2285	2493	2455	2310	2061	2540	2357	2278	2308	2122	1817	1704	1574	1542	1423	1332	965
NT 10-K/A	**627**	0	1	6	21	36	41	36	41	49	43	34	58	37	43	17	27	36	21	21	10	19	8	7	15
NT 10-Q	**63960**	0	68	373	1261	1757	1828	1940	3705	4046	4114	3632	3604	4377	4575	4062	3681	3422	3079	3179	2908	2701	2502	2239	907
NT 10-Q/A	**557**	0	3	7	16	16	18	36	37	38	49	33	53	41	36	32	25	18	15	25	25	12	10	8	4
NT 11-K	**1612**	0	1	36	50	66	106	124	112	75	123	117	143	121	121	124	61	69	25	27	23	22	11	27	28
NT 11-K/A	**15**	0	1	0	1	0	0	2	2	0	0	1	1	1	0	0	0	2	0	0	0	0	0	4	0
NT 15D2	**8**	0	1	1	0	0	1	0	0	1	1	2	0	0	0	1	0	0	0	0	0	0	0	0	0
NT 15D2/A	**7**	0	0	0	1	1	1	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NT 20-F	**1323**	0	0	0	0	5	6	10	18	27	85	120	118	110	117	101	80	82	58	60	81	62	64	64	55
NT 20-F/A	**22**	0	0	0	0	0	0	0	0	1	2	6	1	1	2	1	1	2	0	1	1	0	0	1	2
NT N-MFP	**235**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	29	28	70	20	41	30	17
NT N-MFP1	**11**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
NTFNCSR	**59**	0	0	0	0	0	0	0	0	0	0	0	30	4	3	3	1	3	2	1	6	0	2	3	1
NTFNSAR	**177**	0	2	4	3	14	20	70	3	21	9	4	2	2	1	7	1	3	2	2	3	2	1	1	0
NTN 10D	**5**	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	0	0	0	0	0	1	0
NTN 10K	**1487**	0	2	5	24	58	83	66	92	109	102	111	110	81	80	71	78	122	45	37	46	39	44	57	25
NTN 10Q	**2219**	0	8	9	49	55	135	99	112	259	172	157	124	120	138	130	125	76	58	88	103	67	51	62	22
NTN 11K	**18**	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NTN 20F	**36**	0	0	0	2	0	0	0	0	3	3	3	3	4	4	4	1	3	6	0	0	0	0	0	0
NTN15D2	**1**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NT-NCSR	**326**	0	0	0	0	0	0	0	0	0	0	35	22	30	64	55	28	18	14	12	4	10	5	6	23
NT-NCSR/A	**17**	0	0	0	0	0	0	0	0	0	0	0	1	3	2	0	3	1	2	1	2	1	1	0	0
NT-NSAR	**2562**	1	4	28	100	216	253	333	391	342	144	122	113	159	101	42	61	38	23	16	17	20	11	18	9
NT-NSAR/A	**111**	0	3	0	15	10	5	0	14	3	8	9	3	6	11	4	1	5	2	2	7	0	1	1	1
OIP NTC	**153**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	90	20	15	18	4	2	0
OIP ORDR	**153**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	67	42	13	18	4	2	0
POS 8C	**1585**	0	0	0	9	39	25	16	33	29	25	36	58	48	75	67	82	109	108	81	117	191	144	170	123
POS AM	**37050**	0	163	425	947	1493	1598	1617	1283	1447	1382	1734	2536	2254	2321	2270	2338	2053	2277	2030	1706	1346	1116	1897	817
POS AMC	**1180**	0	144	175	119	110	58	82	121	97	93	87	40	53	1	0	0	0	0	0	0	0	0	0	0
POS AMI	**5399**	1	44	202	320	293	307	233	263	277	274	292	247	232	225	207	244	234	211	216	212	223	227	236	179
POS EX	**3347**	0	0	0	0	9	42	74	75	69	73	79	66	94	166	112	160	121	132	395	490	272	426	266	226
POS462B	**203**	0	0	0	9	33	29	15	18	15	19	5	4	14	4	5	3	5	4	7	5	2	1	3	3
POS462C	**75**	0	0	0	2	8	8	7	10	8	6	5	1	7	2	3	4	0	0	0	0	0	1	3	0
POSASR	**16050**	0	0	0	0	0	0	0	0	0	0	0	0	2	503	779	734	878	1489	694	2725	1500	1983	2455	2308
PRE 14A	**36169**	0	560	774	1636	2017	2109	1934	1931	1549	1871	1775	1843	2003	2043	1930	1821	2046	1552	1212	1100	1277	1249	1161	776
PRE 14C	**8063**	0	4	21	82	133	128	152	254	322	410	466	646	598	613	578	536	498	492	408	346	386	404	417	169
PRE13E3	**223**	0	2	25	34	46	64	43	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PRE13E3/A	**365**	0	11	42	46	77	67	103	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PREA14A	**13**	0	10	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PREA14C	**1**	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PREC14A	**3574**	0	49	61	105	159	169	201	233	188	156	219	105	133	139	181	206	225	145	133	149	157	152	169	140
PREC14C	**78**	0	4	0	0	6	2	1	6	6	5	8	1	3	1	0	6	11	3	3	4	3	4	1	0
PREM14A	**4109**	0	14	37	119	144	161	206	330	298	224	238	225	236	314	343	178	125	183	167	129	128	119	128	63
PREM14C	**604**	0	1	6	14	16	23	44	47	43	48	41	37	32	34	46	35	17	19	25	27	9	17	13	10
PREN14A	**681**	0	2	14	60	51	32	36	28	24	32	25	14	33	66	37	52	55	20	20	18	20	28	10	4
PRER14A	**10217**	0	75	212	375	546	672	566	531	637	577	634	547	559	603	558	569	616	371	307	246	370	263	239	144
PRER14C	**2230**	0	1	9	28	50	47	35	33	70	115	113	137	188	194	146	178	163	147	144	94	103	117	93	25
PRES14A	**4558**	0	64	218	533	775	696	602	795	714	161	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PRES14C	**121**	0	0	2	16	29	12	12	19	26	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PRRN14A	**3678**	0	14	60	141	231	214	215	196	182	158	227	130	161	168	161	134	282	152	133	124	147	152	168	128
PX14A6G	**1805**	0	1	2	0	18	13	13	8	20	30	41	105	55	62	113	48	68	58	98	142	182	188	342	198
PX14A6N	**54**	0	0	0	0	0	0	0	0	6	0	0	2	6	0	10	0	0	0	8	2	0	0	16	4
QRTLYRPT	**68**	0	0	0	0	0	0	0	0	0	0	10	4	3	5	4	3	3	3	3	3	3	7	12	5
QRTLYRPT/A	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
QUALIF	**73**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	58
REGDEX	**126758**	1	0	2	5	2	5	14	23	45	14000	14149	16194	18270	19434	20923	19566	4125	0	0	0	0	0	0	0
REGDEX/A	**41947**	0	0	0	1	0	2	1	3	7	3495	3686	4553	5417	6451	7614	7002	3715	0	0	0	0	0	0	0
REG-NR	**51**	0	0	0	0	0	0	0	0	0	1	8	6	7	3	4	3	1	5	6	0	5	0	2	0
REG-NR/A	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
REVOKED	**4714**	0	0	0	0	0	0	0	0	0	0	1	51	124	154	168	470	495	590	596	437	470	483	466	209
RW	**11642**	0	33	61	214	386	418	403	887	899	698	655	621	701	794	646	650	643	610	581	444	364	356	370	208
RW WD	**360**	0	0	0	0	0	0	0	0	0	18	44	26	29	31	26	41	23	38	28	18	13	14	6	5
S-1	**20279**	0	41	87	916	1093	922	1002	1055	433	482	411	1046	649	1026	976	1352	1198	1359	1472	1115	1119	1297	860	368
S-1/A	**50582**	0	96	191	2457	2860	2395	3261	2944	958	1193	944	2814	1892	2239	2287	2659	2267	3438	3909	3037	2763	3035	2133	810
S-11	**1053**	0	7	0	45	75	69	23	15	29	58	78	92	69	31	29	33	56	70	52	49	70	56	37	10
S-11/A	**3324**	0	9	4	102	237	196	77	29	62	131	185	309	233	118	75	88	199	303	201	191	229	147	125	74
S-11MEF	**78**	0	0	0	4	11	2	0	0	2	2	7	4	11	5	0	2	1	1	4	4	7	9	1	1
S-1MEF	**1229**	0	0	0	100	117	63	88	65	26	35	24	54	46	68	63	17	23	44	37	41	98	110	91	19
S-2	**1069**	0	32	62	119	163	117	74	81	91	72	84	99	75	0	0	0	0	0	0	0	0	0	0	0
S-2/A	**1715**	0	154	116	174	282	173	76	96	115	120	137	142	130	0	0	0	0	0	0	0	0	0	0	0
S-20	**6**	0	0	0	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
S-20/A	**5**	0	0	0	0	2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0
S-2MEF	**80**	0	0	0	11	14	8	0	5	15	9	9	4	5	0	0	0	0	0	0	0	0	0	0	0
S-3	**48169**	1	725	1097	2349	3099	3528	2948	3263	3492	3524	3487	3275	2797	1463	1530	1039	2684	1269	1243	1105	1287	1222	1254	488
S-3/A	**45634**	0	1035	1104	2330	3018	2971	2940	2551	3020	4611	3386	3645	2502	1421	1395	575	2483	1026	1542	693	1492	977	712	205
S-3ASR	**24127**	0	0	0	0	0	0	0	0	0	0	0	0	522	2175	1554	1166	1934	1994	1122	3213	2039	2850	3787	1771
S-3D	**878**	0	41	57	67	93	55	54	40	28	42	38	38	37	22	25	38	39	19	24	30	28	23	24	16

S-3D/A	**5**	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S-3DPOS	**557**	0	36	34	31	53	41	34	17	15	16	20	18	22	20	13	19	31	18	13	33	21	20	21	11		
S-3MEF	**1834**	0	0	11	249	212	156	115	82	244	110	102	70	83	37	69	19	59	42	54	25	43	29	20	3		
S-4	**79564**	0	269	383	1017	2348	3132	3371	1288	4294	3315	5027	5247	4516	3464	3980	1760	1997	5316	6930	4137	6745	4964	4405	1659		
S-4 POS	**327**	0	30	23	26	31	32	26	14	16	11	15	5	17	3	9	6	2	2	2	2	0	4	50	1		
S-4/A	**99270**	0	271	569	1371	3409	5586	5290	2369	4764	5214	7071	7185	6128	4292	4161	2748	2073	7889	8248	6246	7535	3001	2969	881		
S-4EF	**74**	0	2	0	4	17	15	4	9	9	4	3	0	1	1	4	0	0	0	0	0	1	0	0	0		
S-4EF/A	**29**	0	0	0	2	10	5	0	5	1	2	1	0	1	0	2	0	0	0	0	0	0	0	0	0		
S-4MEF	**135**	0	0	0	0	11	25	13	12	14	6	6	7	4	5	5	5	3	0	2	2	2	6	4	3		
S-6	**16101**	0	79	81	69	132	570	634	679	611	537	498	468	525	581	703	789	938	1131	1190	1287	1207	1340	1358	694		
S-6/A	**6202**	0	3	13	9	49	237	288	228	177	153	103	60	75	85	149	221	313	377	373	716	664	748	755	406		
S-6EL24	**1394**	0	312	333	391	358	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
S-6EL24/A	**361**	0	36	89	120	116	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
S-8	**74629**	0	990	1785	4099	5196	5224	5127	5786	5292	4536	4028	4076	3499	3134	2965	2518	2389	2249	2169	1985	2100	2184	2144	1154		
S-8 POS	**25757**	0	131	307	524	730	761	737	771	873	830	882	798	945	1079	1361	1444	1516	2055	1846	1728	1594	1665	2086	1094		
S-8/A	**17**	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
S-B	**291**	0	0	0	0	0	0	0	0	0	35	30	20	16	24	13	19	24	21	13	25	15	12	19	5		
S-B/A	**169**	0	0	0	0	0	0	0	0	0	12	16	14	14	15	8	5	15	16	10	11	5	7	14	7		
SB-1	**150**	0	0	0	2	9	8	17	29	12	15	10	3	4	27	14	0	0	0	0	0	0	0	0	0		
SB-1/A	**354**	0	0	0	7	17	8	31	53	44	30	41	26	18	51	28	0	0	0	0	0	0	0	0	0		
SB-1MEF	**1**	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
SB-2	**7777**	0	0	18	272	398	377	393	738	653	602	434	799	883	1032	1055	123	0	0	0	0	0	0	0	0		
SB-2/A	**16540**	0	0	45	617	935	801	781	1151	1644	1434	1192	1590	2027	2211	1925	187	0	0	0	0	0	0	0	0		
SB-2MEF	**104**	0	0	1	11	9	13	6	2	4	13	9	10	12	5	9	0	0	0	0	0	0	0	0	0		
S-BMEF	**12**	0	0	0	0	0	0	0	0	0	0	0	1	1	1	4	1	0	0	1	0	1	0	2	0		
SC 13D	**100114**	0	613	1373	4523	7527	6695	6305	6242	5505	5215	5403	4908	5112	5538	5499	4920	3778	3878	3691	2951	2895	3033	3011	1499		
SC 13D/A	**201938**	2	1687	3997	7721	10790	10923	10237	10160	9765	9886	10026	10249	9841	10441	11310	10739	9871	9043	8320	7710	8486	7983	8367	4384		
SC 13E1	**21**	0	0	0	5	0	2	2	2	2	1	2	1	0	0	0	0	0	0	1	1	0	0	2	0		
SC 13E1/A	**14**	0	0	0	5	0	2	2	0	1	0	0	0	2	0	0	0	0	1	0	0	0	0	0	1		
SC 13E3	**2629**	0	19	36	91	123	135	244	187	224	137	232	140	160	145	125	82	105	102	94	54	81	43	33	37		
SC 13E3/A	**8213**	0	53	136	292	289	430	609	424	419	375	517	503	633	496	495	344	381	411	386	212	312	258	103	135		
SC 13E4	**752**	0	38	97	127	197	132	152	8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
SC 13E4/A	**1357**	0	56	196	244	349	225	232	53	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
SC 13G	**344030**	0	3145	4691	6591	17494	22365	18645	19321	18256	13333	13330	15159	18007	20018	21207	22123	16883	15011	13153	12823	13321	14023	14854	10277		
SC 13G/A	**587711**	0	5400	9232	13765	20819	24884	25422	24682	23959	25883	25046	24359	24857	28436	32170	33726	33635	29040	28587	29077	30304	31073	32750	30605		
SC 14D1	**2383**	0	68	235	297	467	493	807	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
SC 14D1/A	**10318**	0	456	1281	1382	1838	2483	2493	366	18	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
SC 14D9	**3997**	0	28	128	225	228	261	381	361	255	232	191	164	149	152	277	147	121	127	134	99	90	96	113	38		
SC 14D9/A	**8499**	0	151	288	319	427	457	491	390	451	329	338	277	239	408	498	365	439	566	497	422	295	367	342	143		
SC 14F1	**2168**	0	3	16	32	46	60	81	91	75	119	116	132	171	155	203	157	110	166	130	78	73	79	45	30		
SC 14F1/A	**222**	0	1	0	2	3	5	9	12	3	25	18	21	37	18	22	12	6	10	7	4	0	1	2	4		
SC TO-C	**7227**	0	0	0	0	0	0	0	536	480	391	388	311	337	575	605	534	627	514	428	354	334	281	309	223		
SC TO-I	**7160**	0	0	0	0	0	0	0	163	375	333	348	291	362	434	507	469	567	492	467	541	511	521	523	256		
SC TO-I/A	**12425**	0	0	0	0	0	0	0	297	880	801	884	659	644	760	903	736	1111	838	637	761	740	684	735	355		
SC TO-T	**4459**	0	0	0	0	0	0	0	591	473	387	290	210	217	217	364	234	194	232	210	168	222	188	198	64		
SC TO-T/A	**21590**	0	0	0	0	0	0	0	2023	1889	1796	1475	1213	1147	1236	1768	1067	1254	1436	1253	985	940	866	890	352		
SC13E4F	**15**	0	0	0	0	0	0	0	0	0	1	3	2	3	0	1	1	2	1	0	1	0	0	0	0		
SC13E4F/A	**22**	0	0	0	0	0	0	0	0	0	6	4	4	4	0	2	1	1	0	0	0	0	0	0	0		
SC14D1F	**78**	0	0	0	0	0	0	0	0	0	4	3	4	4	22	16	2	5	2	4	4	4	4	0	0		
SC14D1F/A	**268**	0	0	0	0	0	0	0	0	0	14	12	8	6	74	40	22	1	15	28	4	28	16	0	0		
SC14D9C	**2069**	0	0	0	0	0	0	0	133	120	93	132	52	69	148	210	195	96	144	154	123	124	88	116	72		
SC14D9F	**37**	0	0	0	0	0	0	0	0	0	4	2	3	2	8	7	1	1	1	1	1	2	4	0	0		
SC14D9F/A	**40**	0	0	0	0	0	0	0	0	0	1	2	2	3	15	5	2	0	3	4	1	1	1	0	0		
SD	**3844**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1334	1284	1226		
SD/A	**13**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	2	2		
SDR	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2		
SE	**3486**	0	0	0	0	0	0	0	0	5	588	843	909	748	56	56	33	35	50	39	34	28	33	24	5		
SF-1	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1		
SF-1/A	**7**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7		
SF-3	**86**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	70	16		
SF-3/A	**177**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	102	75		
SL	**9**	0	0	0	0	0	0	0	0	0	5	0	1	0	0	0	0	0	0	0	2	1	0	0	0		
SP 15D2	**55**	0	0	0	2	9	2	5	2	2	0	7	4	1	2	4	2	2	0	1	4	2	3	1	0		
SP 15D2/A	**2**	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0		
STOP ORDER	**42**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	27	11	2		
SUPPL	**55389**	0	0	0	0	0	0	1	0	14	7023	7362	7103	7548	8302	7907	5408	2381	1423	153	231	168	169	133	63		
T-3	**1746**	0	13	11	19	21	27	22	16	80	85	141	46	27	17	106	185	124	173	42	122	142	44	197	86		
T-3/A	**2468**	0	10	9	6	15	31	12	26	54	147	69	64	42	13	92	244	227	261	48	170	196	37	600	95		
TA-1	**323**	0	0	0	0	0	0	0	0	0	45	26	46	42	31	26	26	23	14	12	9	9	3	6	5		
TA-1/A	**3474**	0	0	0	0	0	0	0	0	0	225	247	224	289	282	497	202	274	219	183	153	171	203	214	91		
TA-2	**7423**	0	0	0	0	0	0	0	0	4	866	770	703	617	585	502	471	477	427	402	362	340	316	300	281		
TA-2/A	**623**	0	0	0	0	0	0	0	0	4	46	34	31	33	24	49	58	67	56	34	48	25	44	40	30		
TACO	**90**	0	0	0	0	0	0	0	0	0	4	0	1	0	1	43	0	1	0	37	1	1	1	0	0		
TA-W	**431**	0	0	0	0	0	0	0	0	0	47	41	36	42	36	48	32	38	22	20	13	16	16	14	10		
TH	**22**	0	0	0	0	0	0	0	0	0	0	0	0	7	7	1	3	0	0	2	2	0	0	0	0		
TTW	**28**	0	0	0	0	0	0	0	0	1	6	1	1	2	0	0	2	0	0	4	3	6	2	0	0		
TTW/A	**8**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	6	1	0	0		
U-1	**955**	0	77	102	91	70	82	79	104	81	65	77	75	52	0	0	0	0	0	0	0	0	0	0	0		
U-1/A	**2308**	0	226	250	201	188	202	192	268	202	152	138	145	141	3	0	0	0	0	0	0	0	0	0	0		
U-12-IA	**42**	0	0	0	0	0	0	8	13	4	2	3	7	5	0	0	0	0	0	0	0	0	0	0	0		
U-12-IA/A	**1**	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
U-12-IB	**2537**	0	30	35	17	21	275	365	298	287	261	302	310	226	110	0	0	0	0	0	0	0	0	0	0		
U-12-IB/A	**4**	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0		
U-13-60	**366**	0	20	26	26	28	30	29	28	31	35	35	40	38	0	0	0	0	0	0	0	0	0	0	0		
U-13-60/A	**17**	0	0	3	0	2	1	1	2	1	5	1	0	1	0	0	0	0	0	0	0	0	0	0	0		
U-13E-1	**3**	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0		
U-33-S	**48**	0	0	0	3	4	4	5	6	5	7	5	6	3	0	0	0	0	0	0	0	0	0	0	0		
U-33-S/A	**1**	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
U-3A-2	**1011**	0	52	73	76	90	104	103	97	87	82	86	80	81	0	0	0	0	0	0	0	0	0	0	0		
U-3A-2/A	**161**	0	22	7	6	12	11	11	6	10	23	23	15	15	0	0	0	0	0	0	0	0	0	0	0		
U-3A3-1	**21**	0	1	1	1	2	3	3	2	2	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0		

U-57	**617**	0	2	12	42	92	77	92	96	74	61	23	28	9	9	0	0	0	0	0	0	0	0	0	0
U-57/A	**240**	0	0	0	3	16	14	42	9	50	54	23	19	10	0	0	0	0	0	0	0	0	0	0	0
U5A	**40**	0	0	0	0	1	3	1	14	6	7	1	4	2	1	0	0	0	0	0	0	0	0	0	0
U5A/A	**4**	0	0	0	0	0	0	0	1	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0
U5B	**32**	0	0	1	0	1	3	0	6	10	5	0	3	3	0	0	0	0	0	0	0	0	0	0	0
U5B/A	**10**	0	0	1	0	0	1	0	0	3	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0
U5S	**248**	0	10	15	16	16	17	20	23	20	26	27	29	29	0	0	0	0	0	0	0	0	0	0	0
U5S/A	**68**	0	3	2	3	3	7	4	5	1	12	9	10	9	0	0	0	0	0	0	0	0	0	0	0
U-6B-2	**1291**	0	39	49	50	52	115	91	91	128	164	177	170	157	8	0	0	0	0	0	0	0	0	0	0
U-6B-2/A	**9**	0	0	0	0	1	2	1	0	0	1	3	1	0	0	0	0	0	0	0	0	0	0	0	0
U-7D	**38**	0	0	2	6	0	3	9	0	2	7	3	3	2	1	0	0	0	0	0	0	0	0	0	0
U-7D/A	**66**	0	0	0	4	4	26	18	6	4	2	1	0	1	0	0	0	0	0	0	0	0	0	0	0
U-9C-3	**636**	0	0	0	0	22	55	63	62	75	80	82	94	103	0	0	0	0	0	0	0	0	0	0	0
U-9C-3/A	**16**	0	0	0	0	0	4	2	1	4	1	1	0	3	0	0	0	0	0	0	0	0	0	0	0
UNDER	**41**	0	0	0	0	0	0	0	7	7	7	4	2	5	1	3	2	0	0	3	0	0	0	0	0
UNDER/A	**1**	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UPLOAD	**159065**	0	0	0	0	0	0	0	0	0	0	0	641	8844	13289	12301	13095	15324	17093	17930	15413	16906	12043	13119	3067
WDL-REQ	**55**	0	0	0	0	0	0	0	0	0	5	8	2	1	4	2	6	8	3	4	1	4	6	1	0
X-17A-5	**71022**	0	0	0	0	0	2	0	0	7	5538	5410	5281	5083	5170	5124	4966	4793	4718	4543	4464	4308	4108	4095	3412
X-17A-5/A	**4936**	0	0	0	0	1	2	0	0	2	337	284	477	434	441	408	356	441	611	279	152	135	92	319	165
Total	**15998058**	35	65461	113651	**202614**	**280130**	**313300**	**321262**	**341962**	**351443**	**448884**	**790772**	**1024211**	**1072291**	**1096624**	**1137485**	**1036832**	**938511**	**979708**	**979632**	**974937**	**990462**	**996692**	**994215**	**546944**

Appendix C: „Annual Report Algorithm“

ID.	Description	Regular Expression	Repl. by
1	Decomposition of *“Complete Submission Text File”*	(?s)<DOCUMENT>.*?</DOCUMENT>	
2	Identification of document (file) types	<TYPE>.*	
3	Removal of graphic files	(?s)<TYPE>GRAPHIC.*?</TEXT>	“ “
4	Removal of MS Excel files	(?s)<TYPE>EXCEL.*?</TEXT>	“ “
5	Removal of PDF files	(?s)<TYPE>PDF.*?</TEXT>	“ “
6	Removal of ZIP files	(?s)<TYPE>ZIP.*?</TEXT>	“ “
7	Removal of cover letter	(?s)<TYPE>COVER.*?</TEXT>	“ “
8	Removal of correspondance between SEC staff and EDGAR participant	(?s)<TYPE>CORRESP.*?</TEXT>	“ “
9	Removal of XBRL instance document	(?s)<TYPE>EX-10[01].INS.*?</TEXT>	“ “
10	Removal of XBRL instance document	(?s)<TYPE>EX-99.SDR [KL].INS.*?</TEXT>	“ “
11	Removal of XBRL taxonomy extension schema document	(?s)<TYPE>EX-10[01].SCH.*?</TEXT>	“ “
12	Removal of XBRL taxonomy extension schema document	(?s)<TYPE>EX-99.SDR [KL].SCH.*?</TEXT>	“ “
13	Removal of XBRL taxonomy extension linkbase document	(?s)<TYPE>EX-10[01].CAL.*?</TEXT>	“ “
14	Removal of XBRL taxonomy extension linkbase document	(?s)<TYPE>EX-99.SDR [KL].CAL.*?</TEXT>	“ “
15	Removal of XBRL taxonomy extension definition linkbase document	(?s)<TYPE>EX-10[01].DEF.*?</TEXT>	“ “
16	Removal of XBRL taxonomy extension definition linkbase document	(?s)<TYPE>EX-99.SDR [KL].LAB.*?</TEXT>	“ “
17	Removal of XBRL taxonomy extension labels linkbase document	(?s)<TYPE>EX-10[01].LAB.*?</TEXT>	“ “
18	Removal of XBRL taxonomy extension labels linkbase document	(?s)<TYPE>EX-99.SDR [KL].LAB.*?</TEXT>	“ “
19	Removal of XBRL taxonomy extension presentation linkbase document	(?s)<TYPE>EX-10[01].PRE.*?</TEXT>	“ “
20	Removal of XBRL taxonomy extension presentation linkbase document	(?s)<TYPE>EX-99.SDR [KL].PRE.*?</TEXT>	“ “
21	Removal of XBRL taxonomy extension reference linkbase document	(?s)<TYPE>EX-10[01].REF.*?</TEXT>	“ “
22	Removal of XBRL documents	(?s)<TYPE>XML.*?</TEXT>	“ “
23	Removal of document type information	<TYPE>.*	“ “
24	Removal of sequence information	<SEQUENCE>.*	“ “
25	Removal of filename	<FILENAME>.*	“ “
26	Removal of description	<DESCRIPTION>.*	“ “
27	Removal of head section (including document title)	(?s)(?i)<Head>.*?</Head>	“ “
28	Removal of table content	(?s)(?i)<Table.*?</Table>	“ “
29	Removal of HTML tags and attributes	(?s)<[^>]*>	“ “
			ASCII (ISO 8859-1/ANSI) character set: 32-126 [http://www.w3schools.com/charsets/ref_html_ascii.asp] [http://www.w3schools.com/charsets/ref_html_8859.asp] [http://www.w3schools.com/charsets/ref_html_ansi.asp] [https://www.sec.gov/info/edgar/edgarfm-vol2-v37.pdf] [http://www.htmlhelp.com/reference/charset/]
30	Decoding normal space	&#32;	“ “
31	Decoding exclamation mark	&#33;	“!”
32	Decoding quotation mark	&#34;	“”””
33	Decoding quotation mark	&quot;	“”””

34	Decoding number sign	&#35;	“‘#”
35	Decoding dollar sign	&#36;	“$”
36	Decoding percent sign	&#37;	“%”
37	Decoding ampersand	&#38;	“&”
38	Decoding ampersand	&amp;	“&”
39	Decoding apostrophe	&#39;	“‘”
40	Decoding left parenthesis	&#40;	“(“
41	Decoding right parenthesis	&#41;	“)”
42	Decoding asterisk	&#42;	“*”
43	Decoding plus sign	&#43;	“+”
44	Decoding comma	&#44;	“ ” ‚
45	Decoding hyphen-minus	&#45;	“-”
46	Decoding full stop	&#46;	“ ” .
47	Decoding solidus	&#47;	“/”
48	Decoding digit zero	&#48;	“0”
49	Decoding digit one	&#49;	“1”
50	Decoding digit two	&#50;	“2”
51	Decoding digit three	&#51;	“3”
52	Decoding digit four	&#52;	“4”
53	Decoding digit five	&#53;	“5”
54	Decoding digit six	&#54;	“6”
55	Decoding digit seven	&#55;	“7”
56	Decoding digit eight	&#56;	“8”
57	Decoding digit nine	&#57;	“9”
58	Decoding colon	&#58;	“:”
59	Decoding semicolon	&#59;	“;” ’
60	Decoding less-than sign	&#60;	“<”
61	Decoding less-than sign	&lt;	“<”
62	Decoding equals sign	&#61;	“=”
63	Decoding greater-than sign	&#62;	“>”
64	Decoding greater-than sign	&gt;	“>”
65	Decoding question mark	&#63;	“?”
66	Decoding commercial at	&#64;	“@”
67	Decoding latin capital letter A	&#65;	“A”
68	Decoding latin capital letter B	&#66;	“B”
69	Decoding latin capital letter C	&#67;	“C”
70	Decoding latin capital letter D	&#68;	“D”
71	Decoding latin capital letter E	&#69;	“E”
72	Decoding latin capital letter F	&#70;	“F”
73	Decoding latin capital letter G	&#71;	“G”
74	Decoding latin capital letter H	&#72;	“H”
75	Decoding latin capital letter I	&#73;	“I”
76	Decoding latin capital letter J	&#74;	“J”

77	Decoding latin capital letter K	&#75;	“K”
78	Decoding latin capital letter L	&#76;	“L”
79	Decoding latin capital letter M	&#77;	“M”
80	Decoding latin capital letter N	&#78;	“N”
81	Decoding latin capital letter O	&#79;	“O”
82	Decoding latin capital letter P	&#80;	“P”
83	Decoding latin capital letter Q	&#81;	“Q”
84	Decoding latin capital letter R	&#82;	“R”
85	Decoding latin capital letter S	&#83;	“S”
86	Decoding latin capital letter T	&#84;	“T”
87	Decoding latin capital letter U	&#85;	“U”
88	Decoding latin capital letter V	&#86;	“V”
89	Decoding latin capital letter W	&#87;	“W”
90	Decoding latin capital letter X	&#88;	“X”
91	Decoding latin capital letter Y	&#89;	“Y”
92	Decoding latin capital letter Z	&#90;	“Z”
93	Decoding left square bracket	&#91;	“[“
94	Decoding reverse solidus	&#92;	“\”
95	Decoding right square bracket	&#93;	“]”
96	Decoding circumflex accent	&#94;	“^”
97	Decoding low line	&#95;	“_”
98	Decoding grave accent	&#96;	“‘”
99	Decoding latin small letter a	&#97;	“a”
100	Decoding latin small letter b	&#98;	“b”
101	Decoding latin small letter c	&#99;	“c”
102	Decoding latin small letter d	&#100;	“d”
103	Decoding latin small letter e	&#101;	“e”
104	Decoding latin small letter f	&#102;	“f”
105	Decoding latin small letter g	&#103;	“g”
106	Decoding latin small letter h	&#104;	“h”
107	Decoding latin small letter i	&#105;	“i”
108	Decoding latin small letter j	&#106;	“j”
109	Decoding latin small letter k	&#107;	“k”
110	Decoding latin small letter l	&#108;	“l”
111	Decoding latin small letter m	&#109;	“m”
112	Decoding latin small letter n	&#110;	“n”
113	Decoding latin small letter o	&#111;	“o”
114	Decoding latin small letter p	&#112;	“p”
115	Decoding latin small letter q	&#113;	“q”
116	Decoding latin small letter r	&#114;	“r”
117	Decoding latin small letter s	&#115;	“s”
118	Decoding latin small letter t	&#116;	“t”
119	Decoding latin small letter u	&#117;	“u”

120	Decoding latin small letter v	&#118;	“v”
121	Decoding latin small letter w	&#119;	“w”
122	Decoding latin small letter x	&#120;	“x”
123	Decoding latin small letter y	&#121;	“y”
124	Decoding latin small letter z	&#122;	“z”
125	Decoding left curly bracket	&#123;	“{“
126	Decoding vertical line	&#124;	“|”
127	Decoding right curly bracket	&#125;	“}”
128	Decoding tilde	&#126;	“~”
			ANSI (Windows-1252) character set: 128-159 [http://www.w3schools.com/charsets/ref_html_ansi.asp] [https://www.sec.gov/info/edgar/edgarfm-vol2-v37.pdf] [http://www.htmlhelp.com/reference/charset/]
129	Decoding euro sign	&#128;	“[Euro]”
130	Decoding euro sign	&euro;	“[Euro]”
131	Decoding single low-9 quotation mark	&#130;	“‘“
132	Decoding single low-9 quotation mark	&sbquo;	“‘“
133	Decoding latin small letter f with hook	&#131;	“ “
134	Decoding latin small letter f with hook	&fnof;	“ “
135	Decoding double low-9 quotation mark	&#132;	“”“
136	Decoding double low-9 quotation mark	&bdquo;	“”“
137	Decoding horizontal ellipsis	&#133;	“…“
138	Decoding horizontal ellipsis	&hellip;	“…“
139	Decoding dagger	&#134;	“ “
140	Decoding dagger	&dagger;	“ “
141	Decoding double dagger	&#135;	“ “
142	Decoding double dagger	&Dagger;	“ “
143	Decoding modifier letter circumflex accent	&#136;	“ “
144	Decoding modifier letter circumflex accent	&circ;	“ “
145	Decoding per mille sign	&#137;	“ “
146	Decoding per mille sign	&permil;	“ “
147	Decoding latin capital letter S with caron	&#138;	“S”
148	Decoding latin capital letter S with caron	&Scaron;	“S”
149	Decoding single left-pointing angle quotation mark	&#139;	“‘“
150	Decoding single left-pointing angle quotation mark	&lsaquo;	“‘“
151	Decoding latin capital ligature OE	&#140;	“OE”
152	Decoding latin capital ligature OE	&OElig;	“OE”
153	Decoding Latin capital letter Z with caron	&#142;	“Z”
154	Decoding Latin capital letter Z with caron	&Zcaron;	“Z”
155	Decoding left single quotation mark	&#145;	“‘“
156	Decoding left single quotation mark	&lsquo;	“‘“
157	Decoding right single quotation mark	&#146;	“‘“

158	Decoding right single quotation mark	’	“’“
159	Decoding left double quotation mark	“	“”“
160	Decoding left double quotation mark	“	“”“
161	Decoding right double quotation mark	”	“”“
162	Decoding right double quotation mark	”	“”””
163	Decoding bullet	•	“ “
164	Decoding bullet	•	“ “
165	Decoding en dash	–	“-“
166	Decoding en dash	–	“-“
167	Decoding em dash	—	“-“
168	Decoding em dash	—	“-“
169	Decoding small tilde	˜	“~“
170	Decoding small tilde	˜	“~“
171	Decoding trade mark sign	™	“[trade mark sign]”
172	Decoding trade mark sign	™	“[trade mark sign]”
173	Decoding latin small letter s with caron	š	“s”
174	Decoding latin small letter s with caron	š	“s”
175	Decoding single right-pointing angle quotation mark	›	“’“
176	Decoding single right-pointing angle quotation mark	›	“’“
177	Decoding latin small ligature oe	œ	“oe”
178	Decoding latin small ligature oe	œ	“oe”
179	Decoding latin small letter z with caron	ž	“z”
180	Decoding latin small letter z with caron	ž	“z”
181	Decoding latin capital letter Y with diaeresis	Ÿ	“Y”
182	Decoding latin capital letter Y with diaeresis	Ÿ	“Y”
			ISO 8859-1 (Latin 1) character set: 160-255 [https://www.w3.org/TR/html4/sgml/entities.html] [http://www.w3schools.com/charsets/ref_html_8859.asp] [http://www.w3schools.com/charsets/ref_html_ansi.asp] [https://www.sec.gov/info/edgar/edgarfm-vol2-v37.pdf] [http://www.htmlhelp.com/reference/charset/]
183	Decoding non-breaking space		“ “
184	Decoding non-breaking space		“ “
185	Decoding inverted exclamation	¡	“¡”
186	Decoding inverted exclamation	¡	“¡”
187	Decoding cent sign	¢	“[Cent]”
188	Decoding cent sign	¢	“[Cent]”
189	Decoding pound sign	£	“[Pound Sterling]”
190	Decoding pound sign	£	“[Pound Sterling]”
191	Decoding currency sign	¤	“[Currency]”
192	Decoding currency sign	¤	“[Currency]”
193	Decoding yen sign	¥	“[Yen]”

194	Decoding yen sign	¥	“[Yen]”
195	Decoding broken bar	¦	“ “
196	Decoding broken bar	¦	“ “
197	Decoding section sign	§	“§”
198	Decoding section sign	§	“§”
199	Decoding diaeresis	¨	“ “
200	Decoding diaeresis	¨	“ “
201	Decoding diaeresis	¨	“ “
202	Decoding copyright sign	©	“[copyright sign]”
203	Decoding copyright sign	©	“[copyright sign]”
204	Decoding feminine ordinal indicator	ª	“ “
205	Decoding feminine ordinal indicator	ª	“ “
206	Decoding left-pointing double angle quotation mark	«	“”“
207	Decoding left-pointing double angle quotation mark	«	“”“
208	Decoding not sign	¬	“ “
209	Decoding not sign	¬	“ “
210	Decoding soft hyphen		“ “
211	Decoding soft hyphen		“ “
212	Decoding registered sign	®	“[registered trademark sign]”
213	Decoding registered sign	®	“[registered trademark sign]”
214	Decoding macron	¯	“ “
215	Decoding macron	¯	“ “
216	Decoding degree sign	°	“ “
217	Decoding degree sign	°	“ “
218	Decoding plus-minus sign	±	“[+/-]”
219	Decoding plus-minus sign	±	“[+/-]”
220	Decoding superscript two	²	“ “
221	Decoding superscript two	²	“ “
222	Decoding superscript three	³	“ “
223	Decoding superscript three	³	“ “
224	Decoding acute accent	´	“’”
225	Decoding acute accent	´	“’”
226	Decoding micro sign	µ	“ “
227	Decoding micro sign	µ	“ “
228	Decoding pilcrow sign	¶	“ “
229	Decoding pilcrow sign	¶	“ “
230	Decoding middle dot	·	“ “
231	Decoding middle dot	·	“ “
232	Decoding cedilla	¸	“ ” ¸
233	Decoding cedilla	¸	“ ” ¸
234	Decoding superscript one	¹	“ “
235	Decoding superscript one	¹	“ “
236	Decoding masculine ordinal	º	“ “

237	Decoding masculine ordinal	º	“ “
238	Decoding right-pointing double angle quotation mark	»	“”“
239	Decoding right-pointing double angle quotation mark	»	“”“
240	Decoding vulgar fraction one quarter	¼	“1/4”
241	Decoding vulgar fraction one quarter	¼	“1/4”
242	Decoding vulgar fraction one half	½	“1/2”
243	Decoding vulgar fraction one half	½	“1/2”
244	Decoding vulgar fraction one half	½	“1/2”
245	Decoding vulgar fraction three quarters	¾	“3/4”
246	Decoding vulgar fraction three quarters	¾	“3/4”
247	Decoding inverted question mark	¿	“?”
248	Decoding inverted question mark	¿	“?”
249	Decoding latin capital letter A with grave	À	“A”
250	Decoding latin capital letter A with grave	&#Agrave	“A”
251	Decoding latin capital letter A with acute	Á	“A”
252	Decoding latin capital letter A with acute	Á	“A”
253	Decoding latin capital letter A with circumflex	Â	“A”
254	Decoding latin capital letter A with circumflex	Â	“A”
255	Decoding latin capital letter A with tilde	Ã	“A”
256	Decoding latin capital letter A with tilde	Ã	“A”
257	Decoding latin capital letter A with diaeresis	Ä	“A”
258	Decoding latin capital letter A with diaeresis	Ä	“A”
259	Decoding latin capital letter A with ring above	Å	“A”
260	Decoding latin capital letter A with ring above	Å	“A”
261	Decoding latin capital letter AE	Æ	“AE”
262	Decoding latin capital letter AE	Æ	“AE”
263	Decoding latin capital letter C with cedilla	Ç	“C”
264	Decoding latin capital letter C with cedilla	Ç	“C”
265	Decoding latin capital letter E with grave	È	“E”
266	Decoding latin capital letter E with grave	È	“E”
267	Decoding latin capital letter E with acute	É	“E”
268	Decoding latin capital letter E with acute	É	“E”
269	Decoding latin capital letter E with circumflex	Ê	“E”
270	Decoding latin capital letter E with circumflex	Ê	“E”
271	Decoding latin capital letter E with diaeresis	Ë	“E”
272	Decoding latin capital letter E with diaeresis	Ë	“E”
273	Decoding latin capital letter I with grave	Ì	“I”
274	Decoding latin capital letter I with grave	Ì	“I”
275	Decoding latin capital letter I with acute	Í	“I”
276	Decoding latin capital letter I with acute	Í	“I”
277	Decoding latin capital letter I with circumflex	Î	“I”
278	Decoding latin capital letter I with circumflex	Î	“I”
279	Decoding latin capital letter I with diaeresis	Ï	“I”

280	Decoding latin capital letter I with diaeresis	Ï	“I”
281	Decoding latin capital letter ETH	Ð	“ETH”
282	Decoding latin capital letter ETH	Ð	“ETH”
283	Decoding latin capital letter N with tilde	Ñ	“N”
284	Decoding latin capital letter N with tilde	Ñ	“N”
285	Decoding latin capital letter O with grave	Ò	“O”
286	Decoding latin capital letter O with grave	Ò	“O”
287	Decoding latin capital letter O with acute	Ó	“O”
288	Decoding latin capital letter O with acute	Ó	“O”
289	Decoding latin capital letter O with circumflex	Ô	“O”
290	Decoding latin capital letter O with circumflex	Ô	“O”
291	Decoding latin capital letter O with tilde	Õ	“O”
292	Decoding latin capital letter O with tilde	Õ	“O”
293	Decoding latin capital letter O with diaeresis	Ö	“O”
294	Decoding latin capital letter O with diaeresis	Ö	“O”
295	Decoding multiplication sign	×	“*”
296	Decoding multiplication sign	×	“*”
297	Decoding latin capital letter O with stroke	Ø	“O”
298	Decoding latin capital letter O with stroke	Ø	“O”
299	Decoding latin capital letter U with grave	Ù	“U”
300	Decoding latin capital letter U with grave	Ù	“U”
301	Decoding latin capital letter U with acute	Ú	“U”
302	Decoding latin capital letter U with acute	Ú	“U”
303	Decoding latin capital letter U with circumflex	Û	“U”
304	Decoding latin capital letter U with circumflex	Û	“U”
305	Decoding latin capital letter U with diaeresis	Ü	“U”
306	Decoding latin capital letter U with diaeresis	Ü	“U”
307	Decoding latin capital letter Y with acute	Ý	“Y”
308	Decoding latin capital letter Y with acute	Ý	“Y”
309	Decoding latin capital letter THORN	Þ	“THORN”
310	Decoding latin capital letter THORN	Þ	“THORN”
311	Decoding latin small letter sharp s	ß	“ß”
312	Decoding latin small letter sharp s	ß	“ß”
313	Decoding latin small letter a with grave	à	“a”
314	Decoding latin small letter a with grave	à	“a”
315	Decoding latin small letter a with acute	á	“a”
316	Decoding latin small letter a with acute	á	“a”
317	Decoding latin small letter a with circumflex	â	“a”
318	Decoding latin small letter a with circumflex	â	“a”
319	Decoding latin small letter a with tilde	ã	“a”
320	Decoding latin small letter a with tilde	ã	“a”
321	Decoding latin small letter a with diaeresis	ä	“a”
322	Decoding latin small letter a with diaeresis	ä	“a”

323	Decoding latin small letter a with ring above	&#229;	“a”
324	Decoding latin small letter a with ring above	&aring;	“a”
325	Decoding latin small letter ae	&#230;	“ae”
326	Decoding latin small letter ae	&aelig;	“ae”
327	Decoding latin small letter c with cedilla	&#231;	“c”
328	Decoding latin small letter c with cedilla	&ccedil;	“c”
329	Decoding latin small letter e with grave	&#232;	“e”
330	Decoding latin small letter e with grave	&egrave;	“e”
331	Decoding latin small letter e with acute	&#233;	“e”
332	Decoding latin small letter e with acute	&eacute;	“e”
333	Decoding latin small letter e with circumflex	&#234;	“e”
334	Decoding latin small letter e with circumflex	&ecirc;	“e”
335	Decoding latin small letter e with diaeresis	&#235;	“e”
336	Decoding latin small letter e with diaeresis	&euml;	“e”
337	Decoding latin small letter i with grave	&#236;	“i”
338	Decoding latin small letter i with grave	&igrave;	“i”
339	Decoding latin small letter i with acute	&#237;	“i”
340	Decoding latin small letter i with acute	&iacute;	“i”
341	Decoding latin small letter i with circumflex	&#238;	“i”
342	Decoding latin small letter i with circumflex	&icirc;	“i”
343	Decoding latin small letter i with diaeresis	&#239;	“i”
344	Decoding latin small letter i with diaeresis	&iuml;	“i”
345	Decoding latin small letter eth	&#240;	“eth”
346	Decoding latin small letter eth	&eth;	“eth”
347	Decoding latin small letter n with tilde	&#241;	“n”
348	Decoding latin small letter n with tilde	&ntilde;	“n”
349	Decoding latin small letter o with grave	&#242;	“o”
350	Decoding latin small letter o with grave	&ograve;	“o”
351	Decoding latin small letter o with acute	&#243;	“o”
352	Decoding latin small letter o with acute	&oacute;	“o”
353	Decoding latin small letter o with circumflex	&#244;	“o”
354	Decoding latin small letter o with circumflex	&ocirc;	“o”
355	Decoding latin small letter o with tilde	&#245;	“o”
356	Decoding latin small letter o with tilde	&otilde;	“o”
357	Decoding latin small letter o with diaeresis	&#246;	“o”
358	Decoding latin small letter o with diaeresis	&ouml;	“o”
359	Decoding division sign	&#247;	“/”
360	Decoding division sign	&divide;	“/”
361	Decoding latin small letter o with stroke	&#248;	“o”
362	Decoding latin small letter o with stroke	&oslash;	“o”
363	Decoding latin small letter u with grave	&#249;	“u”
364	Decoding latin small letter u with grave	&ugrave;	“u”
365	Decoding latin small letter u with acute	&#250;	“u”

366	Decoding latin small letter u with acute	ú	“u”
367	Decoding latin small letter u with circumflex	û	“u”
368	Decoding latin small letter u with circumflex	û	“u”
369	Decoding latin small letter u with diaeresis	ü	“u”
370	Decoding latin small letter u with diaeresis	ü	“u”
371	Decoding latin small letter y with acute	ý	“y”
372	Decoding latin small letter y with acute	ý	“y”
373	Decoding latin small letter thorn	þ	“thorn”
374	Decoding latin small letter thorn	þ	“thorn”
375	Decoding latin small letter y with diaeresis	ÿ	“y”
376	Decoding latin small letter y with diaeresis	ÿ	“y”
			Mathematical, Greek and Symbolic characters set: https://www.w3.org/TR/html4/sgml/entities.html http://www.w3schools.com/charsets/ref_html_symbols.asp http://www.w3schools.com/charsets/ref_html_entities_4.asp
377	Decoding latin small f with hook	ƒ	“ “
378	Decoding latin small f with hook	ƒ	“ “
379	Decoding greek capital letter alpha	Α	“ “
380	Decoding greek capital letter alpha	Α	“ “
381	Decoding greek capital letter beta	Β	“ “
382	Decoding greek capital letter beta	Β	“ “
383	Decoding greek capital letter gamma	Γ	“ “
384	Decoding greek capital letter gamma	Γ	“ “
385	Decoding greek capital letter delta	Δ	“ “
386	Decoding greek capital letter delta	Δ	“ “
387	Decoding greek capital letter epsilon	Ε	“ “
388	Decoding greek capital letter epsilon	Ε	“ “
389	Decoding greek capital letter zeta	Ζ	“ “
390	Decoding greek capital letter zeta	Ζ	“ “
391	Decoding greek capital letter eta	Η	“ “
392	Decoding greek capital letter eta	Η	“ “
393	Decoding greek capital letter theta	Θ	“ “
394	Decoding greek capital letter theta	Θ	“ “
395	Decoding greek capital letter iota	Ι	“ “
396	Decoding greek capital letter iota	Ι	“ “
397	Decoding greek capital letter kappa	Κ	“ “
398	Decoding greek capital letter kappa	Κ	“ “
399	Decoding greek capital letter lambda	Λ	“ “
400	Decoding greek capital letter lambda	Λ	“ “
401	Decoding greek capital letter mu	Μ	“ “
402	Decoding greek capital letter mu	Μ	“ “
403	Decoding greek capital letter nu	Ν	“ “

404	Decoding greek capital letter nu	Ν	“ “
405	Decoding greek capital letter xi	Ξ	“ “
406	Decoding greek capital letter xi	Ξ	“ “
407	Decoding greek capital letter omicron	Ο	“ “
408	Decoding greek capital letter omicron	Ο	“ “
409	Decoding greek capital letter pi	Π	“ “
410	Decoding greek capital letter pi	Π	“ “
411	Decoding greek capital letter rho	Ρ	“ “
412	Decoding greek capital letter rho	Ρ	“ “
413	Decoding greek capital letter sigma	Σ	“ “
414	Decoding greek capital letter sigma	Σ	“ “
415	Decoding greek capital letter tau	Τ	“ “
416	Decoding greek capital letter tau	Τ	“ “
417	Decoding greek capital letter upsilon	Υ	“ “
418	Decoding greek capital letter upsilon	Υ	“ “
419	Decoding greek capital letter phi	Φ	“ “
420	Decoding greek capital letter phi	Φ	“ “
421	Decoding greek capital letter chi	Χ	“ “
422	Decoding greek capital letter chi	Χ	“ “
423	Decoding greek capital letter psi	Ψ	“ “
424	Decoding greek capital letter psi	Ψ	“ “
425	Decoding greek capital letter omega	Ω	“ “
426	Decoding greek capital letter omega	Ω	“ “
427	Decoding greek small letter alpha	α	“ “
428	Decoding greek small letter alpha	α	“ “
429	Decoding greek small letter beta	β	“ “
430	Decoding greek small letter beta	β	“ “
431	Decoding greek small letter gamma	γ	“ “
432	Decoding greek small letter gamma	γ	“ “
433	Decoding greek small letter delta	δ	“ “
434	Decoding greek small letter delta	δ	“ “
435	Decoding greek small letter epsilon	ε	“ “
436	Decoding greek small letter epsilon	ε	“ “
437	Decoding greek small letter zeta	ζ	“ “
438	Decoding greek small letter zeta	ζ	“ “
439	Decoding greek small letter eta	η	“ “
440	Decoding greek small letter eta	η	“ “
441	Decoding greek small letter theta	θ	“ “
442	Decoding greek small letter theta	θ	“ “
443	Decoding greek small letter iota	ι	“ “
444	Decoding greek small letter iota	ι	“ “
445	Decoding greek small letter kappa	κ	“ “
446	Decoding greek small letter kappa	κ	“ “

447	Decoding greek small letter lambda	&#955;	“ “
448	Decoding greek small letter lambda	&lambda;	“ “
449	Decoding greek small letter mu	&#956;	“ “
450	Decoding greek small letter mu	&mu;	“ “
451	Decoding greek small letter nu	&#957;	“ “
452	Decoding greek small letter nu	&nu;	“ “
453	Decoding greek small letter xi	&#958;	“ “
454	Decoding greek small letter xi	&xi;	“ “
455	Decoding greek small letter omicron	&#959;	“ “
456	Decoding greek small letter omicron	&omicron;	“ “
457	Decoding greek small letter pi	&#960;	“ “
458	Decoding greek small letter pi	&pi;	“ “
459	Decoding greek small letter rho	&#961;	“ “
460	Decoding greek small letter rho	&rho;	“ “
461	Decoding greek small letter final sigma	&#962;	“ “
462	Decoding greek small letter final sigma	&sigmaf;	“ “
463	Decoding greek small letter sigma	&#963;	“ “
464	Decoding greek small letter sigma	&sigma;	“ “
465	Decoding greek small letter tau	&#964;	“ “
466	Decoding greek small letter tau	&tau;	“ “
467	Decoding greek small letter upsilon	&#965;	“ “
468	Decoding greek small letter upsilon	&upsilon;	“ “
469	Decoding greek small letter phi	&#966;	“ “
470	Decoding greek small letter phi	&phi;	“ “
471	Decoding greek small letter chi	&#967;	“ “
472	Decoding greek small letter chi	&chi;	“ “
473	Decoding greek small letter psi	&#968;	“ “
474	Decoding greek small letter psi	&psi;	“ “
475	Decoding greek small letter omega	&#969;	“ “
476	Decoding greek small letter omega	&omega;	“ “
477	Decoding greek small letter theta symbol	&#977;	“ “
478	Decoding greek small letter theta symbol	&thetasym;	“ “
479	Decoding greek upsilon with hook symbol	&#978;	“ “
480	Decoding greek upsilon with hook symbol	&upsih;	“ “
481	Decoding greek pi symbol	&#982;	“ “
482	Decoding greek pi symbol	&piv;	“ “
483	Decoding bullet	&#8226;	“ “
484	Decoding bullet	&bull;	“ “
485	Decoding horizontal ellipsis	&#8230;	“…“
486	Decoding horizontal ellipsis	&hellip;	“…“
487	Decoding prime	&#8242;	“’”
488	Decoding prime	&prime;	“’”
489	Decoding double prime	&#8243;	“””

490	Decoding double prime	″	“””
491	Decoding overline	‾	“ “
492	Decoding overline	‾	“ “
493	Decoding fraction slash	⁄	“/”
494	Decoding fraction slash	⁄	“/”
495	Decoding blackletter capital I	ℑ	“ “
496	Decoding blackletter capital I	ℑ	“ “
497	Decoding script capital P	℘	“ “
498	Decoding script capital P	℘	“ “
499	Decoding blackletter capital	ℜ	“ “
500	Decoding blackletter capital	ℜ	“ “
501	Decoding trade mark sign	™	[trade mark sign]
502	Decoding trade mark sign	™	[trade mark sign]
503	Decoding alef symbol	ℵ	“ “
504	Decoding alef symbol	ℵ	“ “
505	Decoding leftwards arrow	←	“ “
506	Decoding leftwards arrow	←	“ “
507	Decoding upwards arrow	↑	“ “
508	Decoding upwards arrow	↑	“ “
509	Decoding rightwards arrow	→	“ “
510	Decoding rightwards arrow	→	“ “
511	Decoding downwards arrow	↓	“ “
512	Decoding downwards arrow	↓	“ “
513	Decoding left right arrow	↔	“ “
514	Decoding left right arrow	↔	“ “
515	Decoding downwards arrow with corner leftwards	↵	“ “
516	Decoding downwards arrow with corner leftwards	↵	“ “
517	Decoding leftwards double arrow	⇐	“ “
518	Decoding leftwards double arrow	⇐	“ “
519	Decoding upwards double arrow	⇑	“ “
520	Decoding upwards double arrow	⇑	“ “
521	Decoding rightwards double arrow	⇒	“ “
522	Decoding rightwards double arrow	⇒	“ “
523	Decoding downwards double arrow	⇓	“ “
524	Decoding downwards double arrow	⇓	“ “
525	Decoding left right double arrow	⇔	“ “
526	Decoding left right double arrow	⇔	“ “
527	Decoding for all	∀	“ “
528	Decoding for all	∀	“ “
529	Decoding partial differential	∂	“ “
530	Decoding partial differential	∂	“ “
531	Decoding there exists	∃	“ “
532	Decoding there exists	∃	“ “

533	Decoding empty set	&#8709;	“ “
534	Decoding empty set	&empty;	“ “
535	Decoding nabla	&#8711;	“ “
536	Decoding nabla	&nabla;	“ “
537	Decoding element of	&#8712;	“ “
538	Decoding element of	&isin;	“ “
539	Decoding not an element of	&#8713;	“ “
540	Decoding not an element of	&notin;	“ “
541	Decoding contains as member	&#8715;	“ “
542	Decoding contains as member	&ni;	“ “
543	Decoding n-ary product	&#8719;	“ “
544	Decoding n-ary product	&prod;	“ “
545	Decoding n-ary sumation	&#8721;	“ “
546	Decoding n-ary sumation	&sum;	“ “
547	Decoding minus sign	&#8722;	“-“
548	Decoding minus sign	&minus;	“-“
549	Decoding asterisk operator	&#8727;	“*“
550	Decoding asterisk operator	&lowast;	“*“
551	Decoding square root	&#8730;	“ “
552	Decoding square root	&radic;	“ “
553	Decoding proportional to	&#8733;	“ “
554	Decoding proportional to	&prop;	“ “
555	Decoding infinity	&#8734;	“ “
556	Decoding infinity	&infin;	“ “
557	Decoding angle	&#8736;	“ “
558	Decoding angle	&ang;	“ “
559	Decoding logical and	&#8743;	“ “
560	Decoding logical and	&and;	“ “
561	Decoding logical or	&#8744;	“ “
562	Decoding logical or	&or;	“ “
563	Decoding intersection	&#8745;	“ “
564	Decoding intersection	&cap;	“ “
565	Decoding union	&#8746;	“ “
566	Decoding union	&cup;	“ “
567	Decoding integral	&#8747;	“ “
568	Decoding integral	&int;	“ “
569	Decoding therefore	&#8756;	“ “
570	Decoding therefore	&there4;	“ “
571	Decoding tilde operator	&#8764;	“~“
572	Decoding tilde operator	&sim;	“~“
573	Decoding approximately equal to	&#8773;	“ “
574	Decoding approximately equal to	&cong;	“ “
575	Decoding almost equal to	&#8776;	“ “

576	Decoding almost equal to	&asymp;	“ “
577	Decoding not equal to	&#8800;	“ “
578	Decoding not equal to	&ne;	“ “
579	Decoding identical to	&#8801;	“ “
580	Decoding identical to	&equiv;	“ “
581	Decoding less-than or equal to	&#8804;	“ “
582	Decoding less-than or equal to	&le;	“ “
583	Decoding greater-than or equal to	&#8805;	“ “
584	Decoding greater-than or equal to	&ge;	“ “
585	Decoding subset of	&#8834;	“ “
586	Decoding subset of	&sub;	“ “
587	Decoding superset of	&#8835;	“ “
588	Decoding superset of	&sup;	“ “
589	Decoding not a subset of	&#8836;	“ “
590	Decoding not a subset of	&nsub;	“ “
591	Decoding subset of or equal to	&#8838;	“ “
592	Decoding subset of or equal to	&sube;	“ “
593	Decoding superset of or equal to	&#8839;	“ “
594	Decoding superset of or equal to	&supe;	“ “
595	Decoding circled plus	&#8853;	“ “
596	Decoding circled plus	&oplus;	“ “
597	Decoding circled times	&#8855;	“ “
598	Decoding circled times	&otimes;	“ “
599	Decoding up tack	&#8869;	“ “
600	Decoding up tack	&perp;	“ “
601	Decoding dot operator	&#8901;	“ “
602	Decoding dot operator	&sdot;	“ “
603	Decoding left ceiling	&#8968;	“ “
604	Decoding left ceiling	&lceil;	“ “
605	Decoding right ceiling	&#8969;	“ “
606	Decoding right ceiling	&rceil;	“ “
607	Decoding left floor	&#8970;	“ “
608	Decoding left floor	&lfloor;	“ “
609	Decoding right floor	&#8971;	“ “
610	Decoding right floor	&rfloor;	“ “
611	Decoding left-pointing angle bracket	&#9001;	“ “
612	Decoding left-pointing angle bracket	&lang;	“ “
613	Decoding right-pointing angle bracket	&#9002;	“ “
614	Decoding right-pointing angle bracket	&rang;	“ “
615	Decoding lozenge	&#9674;	“ “
616	Decoding lozenge	&loz;	“ “
617	Decoding black spade suit	&#9824;	“ “
618	Decoding black spade suit	&spades;	“ “

619	Decoding black club suit	&#9827;	“ “
620	Decoding black club suit	&clubs;	“ “
621	Decoding black heart suit	&#9829;	“ “
622	Decoding black heart suit	&hearts;	“ “
623	Decoding black diamond suit	&#9830;	“ “
624	Decoding black diamond suit	&diams;	“ “
			Special characters set: [https://www.w3.org/TR/html4/sgml/entities.html]
625	Decoding latin capital ligature OE	&#338;	“OE”
626	Decoding latin capital ligature OE	&OElig;	“OE”
627	Decoding latin small ligature oe	&#339;	“oe”
628	Decoding latin small ligature oe	&oelig;	“oe”
629	Decoding latin capital letter S with caron	&#352;	“S”
630	Decoding latin capital letter S with caron	&Scaron;	“S”
631	Decoding latin small letter s with caron	&#353;	“s”
632	Decoding latin small letter s with caron	&scaron;	“s”
633	Decoding latin capital letter Y with diaeresis	&#376;	“Y”
634	Decoding latin capital letter Y with diaeresis	&Yuml;	“Y”
635	Decoding modifier letter circumflex accent	&#710;	“ “
636	Decoding modifier letter circumflex accent	&circ;	“ “
637	Decoding small tilde	&#732;	“~“
638	Decoding small tilde	&tilde;	“~“
639	Decoding en space		“ “
640	Decoding en space		“ “
641	Decoding em space		“ “
642	Decoding em space		“ “
643	Decoding thin space		“ “
644	Decoding thin space		“ “
645	Decoding zero width non-joiner	&#8204;	“ “
646	Decoding zero width non-joiner	&zwnj;	“ “
647	Decoding zero width joiner	&#8205;	“ “
648	Decoding zero width joiner	&zwj;	“ “
649	Decoding left-to-right mark	&#8206;	“ “
650	Decoding left-to-right mark	&lrm;	“ “
651	Decoding right-to-left mark	&#8207;	“ “
652	Decoding right-to-left mark	&rlm;	“ “
653	Decoding en dash	&#8211;	“-“
654	Decoding en dash	&ndash;	“-“
655	Decoding em dash	&#8212;	“-“
656	Decoding em dash	&mdash;	“-“
657	Decoding left single quotation mark	&#8216;	“‘“
658	Decoding left single quotation mark	&lsquo;	“‘“

659	Decoding right single quotation mark	&#8217;	‘’‘‘
660	Decoding right single quotation mark	&rsquo;	‘’‘‘
661	Decoding single low-9 quotation mark	&#8218;	‘’‘‘
662	Decoding single low-9 quotation mark	&sbquo;	‘’‘‘
663	Decoding left double quotation mark	&#8220;	“”“‘
664	Decoding left double quotation mark	&ldquo;	“”“‘
665	Decoding right double quotation mark	&#8221;	“”””
666	Decoding right double quotation mark	&rdquo;	“”””
667	Decoding double low-9 quotation mark	&#8222;	“”“‘
668	Decoding double low-9 quotation mark	&bdquo;	“”“‘
669	Decoding dagger	&#8224;	“ “
670	Decoding dagger	&dagger;	“ “
671	Decoding double dagger	&#8225;	“ “
672	Decoding double dagger	&Dagger;	“ “
673	Decoding per mille sign	&#8240;	“ “
674	Decoding per mille sign	&permil;	“ “
675	Decoding single left-pointing angle quotation mark	&#8249;	‘’‘‘
676	Decoding single left-pointing angle quotation mark	&lsaquo;	‘’‘‘
677	Decoding single right-pointing angle quotation mark	&#8250;	‘’‘‘
678	Decoding single right-pointing angle quotation mark	&rsaquo;	‘’‘‘
679	Decoding euro sign	&#8364;	“[Euro]”
680	Decoding euro sign	&euro;	“[Euro]”