Mutual Fund Prospectus Risk/Return Summary Data Sets

Contents

1   Overview

The following data sets provide information extracted from EX-101 exhibits submitted to the Commission in a flattened data format to assist users in more easily consuming the data for analysis. The data is sourced from selected information found in the XBRL tagged mutual fund prospectuses submitted by filers to the Commission. Certain additional fields used in the Commission's EDGAR system are also included to help in supporting the use of the data. The information has been taken directly from submissions created by each registrant, and the data is "as filed" by the registrant. The information will be updated quarterly. Data contained in documents filed after 5:30pm EST on the last business day of the quarter will be included in the next quarterly posting.

DISCLAIMER: The Mutual Fund Prospectus Risk/Return Summary Data Sets contain information derived from structured data filed with the Commission by individual registrants as well as Commission-generated filing identifiers. Because the data sets are derived from information provided by individual registrants, we cannot guarantee the accuracy of the data sets. In addition, it is possible inaccuracies or other errors were introduced into the data sets during the process of extracting the data and compiling the data sets. Finally, the data sets do not reflect all available information, including certain metadata associated with Commission filings. The data sets are intended to assist the public in analyzing data contained in Commission filings; however, they are not a substitute for such filings. Investors should review the full Commission filings before making any investment decision.

The data extracted from the XBRL submissions is organized into data sets containing information about submissions, numbers, taxonomy tags, and more. Each data set consists of rows and fields, and is provided as a tab-delimited TXT format file. The data sets are as follows:

- **SUB** – Submission data set; this includes one record for each submission having an XBRL exhibit. The set includes fields of information pertinent to the submission and the filing entity.
- **TAG** – Tag data set; includes defining information about each tag. Information includes tag descriptions (documentation labels), taxonomy version information and other tag attributes.
- **LAB** – Label data set; this contains additional information about each tag as it was used in a specific exhibit.
- **NUM** – Number data set; this includes one row for each distinct amount from each exhibit included in the SUB data set. The Number data set includes, for every exhibit, all line item values.
- **CAL** – Calculation data set; provides information to arithmetically relate tags in a filing.
- **TXT** – Text data set; this is the plain text of all the non-numeric tagged items in the exhibit.

2   Scope

The scope of the data in the mutual fund prospectus risk/return summary data sets consists of:

- Numeric data and non-numeric "plain text" from the risk/return summary
- From Interactive Data exhibits (XBRL) to forms 485BPOS an 497;
- Submitted from 12/20/2010 through the "Data Cutoff Date" inclusive

All data values are "as filed."

3   Organization

Note that this data set represents quarterly and annual uncorrected and "as filed" EDGAR document submissions containing multiple reporting periods (including amendments of prior submissions). Data in this submitted form may contain redundancies, inconsistencies, and discrepancies relative to other publication formats. Each quarterly data set is accompanied by a metadata file conforming to the W3C specification for tabular data (https://www.w3.org/TR/2015/REC-tabular-data-model-20151217/) that encodes the following information about the data sets and their relationships to each other.

1. **SUB** is identifies all the EDGAR submissions in the data set, with each row having the unique (primary) key **adsh,** a 20 character EDGAR Accession Number with dashes in positions 11 and 14.
2. **TAG** is a data set of all tags used in the submissions, both standard and custom. These fields comprise a unique compound key:
   1) **tag –** tag used by the filer
   2) **version** – if a standard tag, the taxonomy of origin, otherwise equal to adsh.
3. **LAB** is a data set of all tag labels appearing in the submission, both standard and custom. These fields comprise a unique compound key:
   1) **adsh –** EDGAR accession number
   2) **tag –** tag used by the filer
   3) **version** – if a standard tag, the taxonomy of origin, otherwise equal to adsh.
4. **CAL** is a data set that provides arithmetic relationships among the tags in a filing. These fields comprise a unique compound key:
   1) **adsh** - EDGAR accession number
   2) **grp** - sequential number of a group of relationships within the submission
   3) **arc -** sequential number of relationship within the group of relationships

5. **NUM** is a data set of all numeric XBRL facts presented on the primary financial statements. These fields comprise a unique compound key:
   1) **adsh** - EDGAR accession number
   2) **tag -** tag used by the filer
   3) **version** – if a standard tag, the taxonomy of origin, otherwise equal to adsh.
   4) **ddate** - document date
   5) **uom -** unit of measure
   6) **series –** the 10-character series identifier, if any
   7) **class –** the 10-character class identifier, if any
   8) **measure –** a token to qualify which performance measure the fact represents, if any
   9) **document** – a token to distinguish between the same fact appearing in different prospectuses within a single submission, if needed
   10) **otherdims** – other tokens with information provided by the filer
   11) **iprx** – a sequential integer used to distinguish otherwise identical facts

6. **TXT** is a data set that contains the plain (no HTML) text of each non-numeric XBRL fact. These fields comprise a unique compound key:
   1) **adsh** - EDGAR accession number
   2) **tag –** tag used by the filer
   3) **version** – if a standard tag, the taxonomy of origin, otherwise equal to adsh
   4) **ddate** - period end date
   5) **lang** – the language (usually US English) of the text
   6) **series –** the 10-character series identifier, if any
   7) **class –** the 10-character class identifier, if any
   8) **measure –** a token to qualify which performance measure the fact represents, if any
   9) **document** – a token to distinguish between the same fact appearing in different prospectuses within a single submission, if needed
   10) **otherdims** – other tokens with information provided by the filer
   11) **iprx** – a sequential integer used to distinguish otherwise identical facts

The relationship of the data sets is as shown in Figure 1. The Accession Number (**adsh)** found in the NUM data set can be used to retrieve information about the submission in SUB. Each row of data in NUM or TXT was tagged by the filer using a tag. Information about the tag used can be found in TAG. Each row of data in NUM or TXT appears on one or more lines of reports detailed in PRE.

Figure 1. Data relationships

Data set	Fields referencing other datasets	Referenced dataset	Referenced fields
NUM	adsh	SUB	adsh
	tag, version	TAG	tag, version
TXT	adsh	SUB	adsh
	tag, version	TAG	tag, version
LAB	adsh	SUB	adsh
	tag, version	TAG	tag, version
CAL	adsh	SUB	adsh
	ptag, pversion	TAG	tag, version
	ctag, cversion	TAG	tag, version

4   File Formats

Each of the data sets is provided in a single encoding, as follows:

> Tab Separated Value (.tsv): utf-8, tab-delimited, \n- terminated lines, with the first line containing the field names in lowercase.

5   Table Definitions

The fields in the figures below provide the following information: field name, description, source, data format, maximum field size, an indication of whether or not the field may be NULL (yes or no), and key.

The Source field has two possible values:

- EDGAR indicates that the source of the data is the filer's EDGAR submission header.
- XBRL indicates that the source of the data is the filer's EX-101 (XBRL) exhibits.

The Key field indicates whether the field is part of a unique index on the data. There are two possible values for this field:

- "*" – Indicates the field is part of the unique key for the row.
- Empty (nothing in field) – the field is not part of the unique compound key.

5.1   SUB (Submissions)

The submissions data set contains summary information about an entire EDGAR submission. Some fields were sourced directly from EDGAR submission information, while other fields of data were sourced from the Interactive Data exhibits of the submission. Note: EDGAR derived fields represent the most recent EDGAR assignment as of a given filing's submission date and do not necessarily represent the most current assignments.

Figure 2. Fields in the SUB data set

Field Name	Field Description	Source	Format	Max Size	May be NULL	Key
adsh	Accession Number. The 20-character string formed from the 18-digit number assigned by the Commission to each EDGAR submission.	EDGAR	ALPHANUMERIC (nnnnnnnnnn-nn-nnnnnn)	20	No	*
cik	Central Index Key (CIK). Ten digit number assigned by the Commission to each registrant that submits filings.	EDGAR	NUMERIC	10	No	
name	Name of registrant. This corresponds to the name of the legal entity as recorded in EDGAR as of the filing date.	EDGAR	ALPHANUMERIC	150	No	
countryba	The ISO 3166-1 country of the registrant's business address.	EDGAR	ALPHANUMERIC	2	No	
stprba	The state or province of the registrant's business address, if field countryba is US or CA.	EDGAR	ALPHANUMERIC	2	Yes	

Field Name	Field Description	Source	Format	Max Size	May be NULL	Key
cityba	The city of the registrant's business address.	EDGAR	ALPHANUMERIC	30	No	
zipba	The zip code of the registrant's business address.	EDGAR	ALPHANUMERIC	10	Yes	
bas1	The first line of the street of the registrant's business address.	EDGAR	ALPHANUMERIC	40	Yes	
bas2	The second line of the street of the registrant's business address.	EDGAR	ALPHANUMERIC	40	Yes	
baph	The phone number of the registrant's business address.	EDGAR	ALPHANUMERIC	12	Yes	
countryma	The ISO 3166-1 country of the registrant's mailing address.	EDGAR	ALPHANUMERIC	2	Yes	
stprma	The state or province of the registrant's mailing address, if field countryma is US or CA.	EDGAR	ALPHANUMERIC	2	Yes	
cityma	The city of the registrant's mailing address.	EDGAR	ALPHANUMERIC	30	Yes	
zipma	The zip code of the registrant's mailing address.	EDGAR	ALPHANUMERIC	10	Yes	
mas1	The first line of the street of the registrant's mailing address.	EDGAR	ALPHANUMERIC	40	Yes	
mas2	The second line of the street of the registrant's mailing address.	EDGAR	ALPHANUMERIC	40	Yes	
countryinc	The country of incorporation for the registrant.	EDGAR	ALPHANUMERIC, ISO 3166-1 ALPHA 2	2	No	
stprinc	The state or province of incorporation for the registrant, if countryinc is US or CA, otherwise NULL.	EDGAR	ALPHANUMERIC	2	Yes	
ein	Employee Identification Number, 9 digit identification number assigned by the Internal Revenue Service to business entities operating in the United States.	EDGAR	NUMERIC	10	Yes	
former	Most recent former name of the registrant, if any.	EDGAR	ALPHANUMERIC	150	Yes	
changed	Date of change from the former name, if any.	EDGAR	ALPHANUMERIC	8	Yes	
fye	Fiscal Year End Date.	XBRL	ALPHANUMERIC (mmdd)	4	No	
pdate	Prospectus date	XBRL	DATE (yyyymmdd)	8	Yes	
effdate	Document Effective Date	XBRL	DATE (yyyymmdd)	8	Yes	
form	The submission type of the registrant's filing.	EDGAR	ALPHANUMERIC	20	No	

Field Name	Field Description	Source	Format	Max Size	May be NULL	Key
filed	The date of the registrant's filing with the Commission.	EDGAR	DATE (yymmdd)	8	No	
accepted	The acceptance date and time of the registrant's filing with the Commission. Filings accepted after 5:30pm EST are considered filed on the following business day.	EDGAR	DATETIME (yyyy-mm-dd hh:mm:ss)	19	No	
instance	The name of the submitted XBRL Instance Document (EX-101.INS) type data file. The name often begins with the company ticker symbol.	EDGAR	ALPHANUMERIC (example: abcd-yyyymmdd.xml)	32	No	
nciks	Number of Central Index Keys (CIK) of registrants (i.e., business units) included in the consolidating entity's submitted filing.	EDGAR	NUMERIC	4	No	
aciks	Additional CIKs of co-registrants included in a consolidating entity's EDGAR submission, separated by spaces. If there are no other co-registrants (i.e., nciks = 1), the value of aciks is NULL. For a very small number of filers, the list of co-registrants is too long to fit in the field. Where this is the case, PARTIAL will appear at the end of the list indicating that not all co-registrants' CIKs are included in the field; users should refer to the complete submission file for all CIK information.	EDGAR	ALPHANUMERIC (space delimited)	120	Yes	

Note: To access the complete submission files for a given filing, please see the Commission EDGAR website. The Commission website folder https://www.sec.gov/Archives/edgar/data/{**cik**}/{**accession**}/ will always contain all the data sets for a given submission. To assemble the folder address to any filing referenced in the SUB data set, simply substitute {**cik**} with the **cik** field and replace {**accession**} with the **adsh** field (after removing the dash character). The following sample SQL Query provides an example of how to generate a list of addresses for filings contained in the SUB data set:

- ```
  select name,form,period, 'https://www.sec.gov/Archives/edgar/data/' +
  ltrim(str(cik,10))+'/' + replace(adsh,'-','')+'/'+instance as url from sub order by
  period desc, name
  ```

5.2 TAG (Tags)

The TAG data set contains all standard taxonomy tags, not just those appearing in submissions to date, and also includes all custom taxonomy tags defined in the submissions. The source is the "as filed" XBRL filer

submissions. The standard tags are derived from taxonomies in
https://www.sec.gov/info/edgar/edgartaxonomies.shtml as of the date of the original submission.

Figure 3. Fields in the TAG data set

Field Name	Field Description	Field Type	Max Size	May be NULL	Key
tag	The unique identifier (name) for a tag in a specific taxonomy release.	ALPHANUMERIC	256	No	*
version	For a standard tag, an identifier for the taxonomy; otherwise the accession number where the tag was defined.	ALPHANUMERIC	20	No	*
custom	1 if tag is custom (version=adsh), 0 if it is standard. *Note: This flag is technically redundant with the version and adsh fields.*	BOOLEAN (1 if true and 0 if false)	1	No	
abstract	1 if the tag is not used to represent a numeric fact.	BOOLEAN (1 if true and 0 if false)	1	No	
datatype	If abstract=1, then NULL, otherwise the data type (e.g., monetary) for the tag.	ALPHANUMERIC	20	Yes	
iord	If abstract=1, then NULL; otherwise, "I" if the value is a point in time, or "D" if the value is a duration.	ALPHANUMERIC	1	No	
tlabel	If a standard tag, then the label text provided by the taxonomy, otherwise the text provided by the filer. A tag which had neither would have a NULL value here.	ALPHANUMERIC	512	Yes	
doc	The detailed definition for the tag, truncated to 2048 characters. If a standard tag, then the text provided by the taxonomy, otherwise the text assigned by the filer. Some tags have neither, in which case this field is NULL.	ALPHANUMERIC	2048	Yes	

5.3   LAB (Labels)

The LAB data set contains all standard taxonomy tags, not just those appearing in submissions to date, and also includes all custom taxonomy tags defined in the submissions. The source is the "as filed" XBRL filer submissions. The standard tags are derived from taxonomies in
https://www.sec.gov/info/edgar/edgartaxonomies.shtml as of the date of the original submission.

Figure 4. Fields in the LAB data set

Field Name	Field Description	Field Type	Max Size	May be NULL	Key
adsh	Accession Number. The 20-character string formed from the 18-digit number assigned by the Commission to each EDGAR submission.	ALPHANUMERIC	20	No	*
tag	The unique identifier (name) for a tag in a specific taxonomy release.	ALPHANUMERIC	256	No	*
version	For a standard tag, an identifier for the taxonomy; otherwise the accession number where the tag was defined.	ALPHANUMERIC	20	No	*
std	Standard label as provided in the submission	ALPHANUMERIC	256	No	
terse	Terse label as provided in the submission	ALPHANUMERIC	256	No	

Field Name	Field Description	Field Type	Max Size	May be NULL	Key
verbose	Verbose label as provided in the submission	ALPHANUMERIC	256	No	
total	Total label as provided in the submission		256		
negated	Negated label as provided in the submission	ALPHANUMERIC	256	No	
negated Terse	Negated, Terse label as provided in the submission	ALPHANUMERIC	256	No	

5.4 CAL (Calculations)

The CAL data set contains one row for each calculation relationship ("arc"). Note that XBRL allows a parent element to have more than one distinct set of arcs for a given parent element, thus the rationale for distinct fields for the group and the arc. The source for the table is the "as filed" XBRL filer submissions.

Figure 5. Fields in the CAL data set

Field Name	Field Description	Field Type (format)	Max Size	May be NULL	Key
adsh	Accession Number. The 20-character string formed from the 18-digit number assigned by the Commission to each EDGAR submission.	ALPHANUMERIC	20	No	*
grp	Sequential number for grouping arcs in a submission.	NUMERIC(1,0)	1	No	*
arc	Sequential number for arcs within a group in a submission.	NUMERIC(1,0)	1	No	*
negative	Indicates a weight of -1 (TRUE if the arc is negative), but typically +1 (FALSE).	BOOLEAN	1	No	
ptag	The tag for the parent of the arc	ALPHANUMERIC	256	No	
pversion	The version of the tag for the parent of the arc	ALPHANUMERIC	20	No	
ctag	The tag for the child of the arc	ALPHANUMERIC	256	No	
cversion	The version of the tag for the child of the arc	ALPHANUMERIC	20	No	

5.5 NUM (Numbers)

The NUM data set contains numeric data, one row per data point in the financial statements. The source for the table is the "as filed" XBRL filer submissions.

Figure 6. Fields in the NUM data set

Field Name	Field Description	Field Type (format)	Max Size	May be NULL	Key
adsh	Accession Number. The 20-character string formed from the 18-digit number assigned by the Commission to each EDGAR submission.	ALPHANUMERIC	20	No	*
tag	The unique identifier (name) for a tag in a specific taxonomy release.	ALPHANUMERIC	256	No	*
version	For a standard tag, an identifier for the taxonomy; otherwise the accession number where the tag was defined.	ALPHANUMERIC	20	No	*
ddate	The end date for the data value, rounded to the nearest month end.	DATE (yyyymmdd)	8	No	*

Field Name	Field Description	Field Type (format)	Max Size	May be NULL	Key
uom	The unit of measure for the value.	ALPHANUMERIC	20	No	*
series	Series identifier to which the fact applies	ALPHANUMERIC	10	No	*
class	Series identifier to which the fact applies	ALPHANUMERIC	10	No	*
measure	Performance measure qualifier for the fact	ALPHANUMERIC	256	No	*
document	Prospectus to which the fact applies	ALPHANUMERIC	256	No	*
otherdims	Other qualifiers provided by the filer to distinguish otherwise identical facts.	ALPHANUMERIC	256	No	*
iprx	A positive integer to distinguish different reported facts that otherwise would have the same primary key. For most purposes, data with iprx greater than 1 are not needed. The priority for the fact is based on higher precision.	NUMERIC	2	No	*
value	The value. This is not scaled, it is as found in the Interactive Data file, but is rounded to four digits to the right of the decimal point.	NUMERIC	16	Yes	
footnote	The plain text of any superscripted footnotes on the value, if any, as shown on the statement page, truncated to 512 characters.	ALPHANUMERIC	512	Yes	
footlen	Number of bytes in the plain text of the footnote prior to truncation; zero if no footnote.	NUMERIC	4	No	
dimn	Small integer representing the number of dimensions. Note that this value is a function of the dimension segments.	NUMERIC	1	No	
dcml	The value of the fact "decimals" attribute, with INF represented by 32767.	NUMERIC	2	No	

5.6   TXT (Plain Text)

The TXT data set contains non-numeric data, one row per data point in the financial statements. The source for the table is the "as filed" XBRL filer submissions.

Figure 7. Fields in the TXT data set

Field Name	Field Description	Field Type (format)	Max Size	May be NULL	Key
adsh	Accession Number. The 20-character string formed from the 18-digit number assigned by the Commission to each EDGAR submission.	ALPHANUMERIC	20	No	*
tag	The unique identifier (name) for a tag in a specific taxonomy release.	ALPHANUMERIC	256	No	*
version	For a standard tag, an identifier for the taxonomy; otherwise the accession number where the tag was defined. For example, "invest/2013" indicates that the tag is defined in the 2013 INVEST taxonomy.	ALPHANUMERIC	20	No	*
ddate	The end date for the data value, rounded to the nearest month end.	DATE (yyyymmdd)	8	No	*
lang	The ISO language code of the fact content.	ALPHANUMERIC	5	No	
series	Series identifier to which the fact applies	ALPHANUMERIC	10	No	*

Field Name	Field Description	Field Type (format)	Max Size	May be NULL	Key
class	Series identifier to which the fact applies	ALPHANUMERIC	10	No	*
measure	Performance measure qualifier for the fact	ALPHANUMERIC	256	No	*
document	Prospectus to which the fact applies	ALPHANUMERIC	256	No	*
otherdims	Other qualifiers provided by the filer to distinguish otherwise identical facts.	ALPHANUMERIC	256	No	*
iprx	A positive integer to distinguish different reported facts that otherwise would have the same primary key. For most purposes, data with iprx greater than 1 are not needed. The priority for the fact based on higher precision.	NUMERIC	2	No	*
dcml	The value of the fact "xml:lang" attribute, en-US represented by 32767, other "en" dialects having lower values, and other languages lower still.	NUMERIC	2	No	
dimn	Small integer representing the number of dimensions, useful for sorting. Note that this value is function of the dimension segments.	NUMERIC(1,0)	1	Yes	
escaped	Flag indicating whether the value has had HTML tags removed.	BOOLEAN	1	No	
srclen	Number of bytes in the original, unprocessed value. Zero indicates a NULL value.	NUMERIC(4,0)	4	No	
txtlen	The original length of the whitespace normalized value, which may have been greater than 2048.	NUMERIC(4,0)	4	Yes	
footnote	The plain text of any superscripted footnotes on the value, as shown on the page, truncated to 512 characters, or if there is no footnote, then this field will be blank.	ALPHANUMERIC	512	Yes	
footlen	Number of bytes in the plain text of the footnote prior to truncation.	NUMERIC(4,0)	4	Yes	
context	The value of the contextRef attribute in the source XBRL document, which can be used to recover the original HTML tagging if desired.	ALPHANUMERIC(255)	255	No	
value	The value, with all whitespace normalized, that is, all sequences of line feeds, carriage returns, tabs, non-breaking spaces, and spaces having been collapsed to a single space, and no leading or trailing spaces. Escaped XML that appears in EDGAR "Text Block" tags is processed to remove all mark-up (comments, processing instructions, elements, attributes). The value is truncated to a maximum number of bytes. The resulting text is not intended for end user display but only for text analysis applications.	ALPHANUMERIC	2048	Yes	