AfricaAI: Runbook - CHEF – COMPENSWISS

Summary	We are collecting data from the Compenswiss fund's annual report
Source name	Compenswiss - Fonds de compensation AVS - COMPENSWISS
Link to the source	https://ar.compenswiss.ch/en_GB/investments https://ar.compenswiss.ch/en_GB/investments/strategic-asset-allocation-sva
Dataset name	Switzerland (CHE) Funds - CHEF
Frequency	Annual
SLA	Exact dates of update of reports on the sources are not specified. The estimated period for updating annual reports is usually the beginning of April. Start checking for new releases every week from next January. Since the beginning of March, the probability of updating reports increases, so checking for releases should be done daily. Since the beginning of April, most of the releases are expected, so you need to check the updates on the resources very carefully every day.
Expected Time	NA

Do not contact the source. Escalate to Data Ops team with any issues in runbook execution that can't be completed with standard plays.

Instructions:

1. Navigate to the source URL.
   https://ar.compenswiss.ch/en_GB/investments
2. Navigate to the "Detailed investment performance" table (scraping asset levels). For each line-item in this table, produce one annual DF series capturing the "Amount" column.

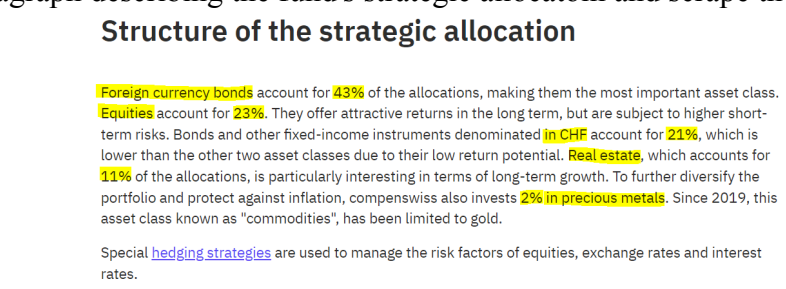
Detailed investment performance

Category	Amount (in CHF million) [1]	Performance (Portfolio in %)
Money market investments	**916**	**-0.37**
Loans	**1 328**	**-1.17**
Swiss francs bonds	**4 717**	**-2.12**
Foreign currency bonds:	**14 179**	**0.04**
Government bonds	3 985	-3.31
Inflation-protected bonds	1 569	6.47
Emerging market bonds	1 267	-2.39
Corporate bonds	3 895	-1.05
High yield bonds	2 004	5.73
Securitised bonds	1 460	3.43
Equities:	**9 656**	**16.69**
Large caps	5 999	21.94
Emerging markets	2 022	-0.72
Small and mid caps	1 636	20.05
Real estate:	**5 085**	**15.69**
Switzerland direct	176	10.08
Switzerland listed	2 083	7.51
Global listed	948	46.68
Global not listed	1 878	10.54
Gold	**1 135**	**-0.1**
Multi-Asset portfolio	**1 002**	**3.36**
Market portfolio	**38 018**	**5.61**
Hedging of currency risk	266	-1.43
Hedging of interest rate risk 4[2]	0	2.89
Equitiy Overlay		0
Market portfolio after hedging	**38 284**	**5.28** [3]
Basic portfolio - Treasury	**2 598**	**-0.34** [4]

Category	**Asset**	Amount in CHF million
Money market investment	MoneyMarket	916
Loans	Loans	1328
Swiss francs bonds	DomesticBondsLocalFX	4717
Foreign currency bonds	ForeignBonds	14179
Government bonds	GovernmentBonds	3985
Inflation-protected Bonds	ILBonds	1569
Emerging markets Bonds	EMBonds	1267
Corporate bonds	CorporateBonds	3895
Higher Yielding Bonds	HighYield	2004

Higher Yielding Bonds	OtherBonds	1460
Equities	Equities	9656
Large Caps	LargeCapEquities	5999
Emerging markets Bonds	EMEquities	2022
Small and mid caps	SmallCapEquities	1636
Real Estate	RealEstate	5085
Swiss Direct	DomesticRealEstate	176
Swiss Listed	DomesticRealEstate	2083
Global Listed	ForeignRealEstate	948
Global not listed	ForeignRealEstate	1878
Gold	Gold	1135
Multi-Asset	MultiAsset	1002
Market Portfolio	Total	38018
Currency Hedging	CurrencyHedging	266
Interest Rate Hedging	InterestRateHedging	0
Equity Overlay	EquityOverlay	NA
Market Portfolio After Hedging	Total	38284
Basic Portfolio - Treasury	Cash	2598

3. Now, navigate to the next URL (strategic allocation) https://ar.compenswiss.ch/en_GB/investments/strategic-asset-allocation-sva
4. Read the paragraph describing the fund's strategic allocatoin and scrape the values accordingly:



Structure of the strategic allocation

Foreign currency bonds account for 43% of the allocations, making them the most important asset class. Equities account for 23%. They offer attractive returns in the long term, but are subject to higher short-term risks. Bonds and other fixed-income instruments denominated in CHF account for 21%, which is lower than the other two asset classes due to their low return potential. Real estate, which accounts for 11% of the allocations, is particularly interesting in terms of long-term growth. To further diversify the portfolio and protect against inflation, compenswiss also invests 2% in precious metals. Since 2019, this asset class known as "commodities", has been limited to gold.

Special hedging strategies are used to manage the risk factors of equities, exchange rates and interest rates.

Structure of the stategic allocaion	**Asset**	**%**
Foreign Currency Bonds	ForeignBonds	43.0
Equities	Equities	23
Bonds in CHF	DomesticBondsLocalFX	21
Real Estate	RealEstate	11
Precious Metals	PreciousMetals	2

7. Save data in CHEF_COMPENSWISS_DATA_YYYYMMDD file.

Requirements to the data provided by AfricaAI

Data files format

For on-boarding new datasets and processing new releases the following requirements are applied:

- there should be one archive in .ZIP format that contains 2 files provided in .XLS format per one release - one file with DATA and one file with METADATA (both files are mandatory). Archive name should be as "**DATASET_YYYYMMDD.ZIP**"", where "DATASET" - specified abbreviation for a dataset, "YYYYMMDD" - timestamp when the file was created, where YYYY - year, MM - month number, DD – day;
- the files format is strictly defined and cannot be changed by AfricaAI team.

File with DATA

Template for file name: “DATASET_DATA_YYYYMMDD.xls”, i.e.” MEXPENFND_DATA_20220601.xls”

Requirements to the file with DATA:

1. format - XLS, encoding - Unicode (UTF-8);
2. file name: should follow pattern "**DATASET_DATA_YYYYMMDD**.xls", where "DATASET" - specified abbreviation for a dataset, "YYYYMMDD" - timestamp when the file was created, where YYYY - year, MM - month number, DD - day;
3. file content:

a. for data with daily/ weekdaily frequency: **YYYY**-**MM**-**DD**, where YYYY - year, MM - month, DD – day;
b. for data with weekly frequency: **YYYY**-**WW**, where YYYY - year, WW - week number (should be two-digits, i.e. 2022-01, 2022-02, etc.). Week number is defined based on the reporting date: reporting date is considered the end of the week, i.e. if reporting date is Wednesday the week is calculated from Thursday till Wednesday – based on that rule the number of the week is defined;
c. for data with monthly frequency: **YYYY**-**MM**, MM - month number (should be two-digits, i.e. 2022-01, 2022-02, etc.);
d. for data with quarterly frequency: **YYYY**-Q**X**, where YYYY - year, X - quarter number;
e. for data with annual frequency: **YYYY**, where YYYY – year;
f. if provider specifies for reported period that data is Not Available (that could provided in corresponding notes or other way) - fill in that data point with "NA" value;
g. if provider does not specify that data is Not Available and report Empty value for reported period - leave Empty cell;
h. if provider skips some of reporting period - skip it also in the file with data as well;
i. data is provided on one sheet only (multiple sheets are forbidden);
j. two-dimensional format where rows represent reported periods and columns contain timeseries values;
k. reported period(s) should be provided in the following format:
l. values to be provided in decimal format with "." as delimiter (number of digits after "." is limited only by provided value by the source). In order to interpret data correctly the following requirements are applied:
m. column names to be provided in two first rows: in the first row - timeseries ID (CODE), in the second row – timeseries description (description in DATA file is identical to description in file with METADATA).

NB! Timeseries with different frequencies could be provided within the same file as there are different patterns for providing reported periods

File with METADATA

Template for file name: “DATASET_META_YYYYMMDD.xls”, i.e.” MEXPENFND_META_20220601.xls”

Requirements to the file with METADATA:

1. format - XLS, encoding - Unicode (UTF-8);
2. file name: should follow pattern "**DATASET_META_YYYYMMDD**.xls", where "DATASET" - specified abbreviation for a dataset, "YYYYMMDD" - timestamp when the file was created, where YYYY - year, MM - month number, DD - day;
3. file content: specified metadata attributes that are static unless provider changes something or new timeseries should be added. In case of new timeseries should be added/ updated/ changed - escalate that to Data Engineering team Special BWAT-DFAC Data Engineering <SpecialBWAT-DFACDataEngineering@epam.com> at and Cc: Data Operations team at Special BWAT-DFAC DataOps <SpecialBWAT-DFACDataOps@epam.com>

Notes *on METADATA fields from file “DATASET_META_YYYYMMDD.xls”*

1. *With every data release attach file with METADATA*
2. *With every release change only on attribute in the file with METADATA in column “NEXT_RELEASE_DATE” – set “YYYY-MM_DDThh:mm:ss” (YYYY – year, MM – month, DD – day, hh – hour, mm – minutes, ss – seconds. Time should be provide in UTC time) to the time when the next release is going to be processed.*

Saving ZIP archive on SFTP

Save the releases in specified per “DATASET” folder (you may create new folder for new DATASET in case of absence of such folder).

Details on SFTP are the following:

1) Dev env

Host: africaai.datadev.dfac.io

2) PRODVAL env

Host: africaai.prodval.data-factory.io

3) PRODHA env

Host: africaai.prodha.data-factory.io
Port: 22