Coursera Project

IBM Applied Data Science Capstone

iShares ETFs World Map with Cost-Return Analysis

By: Nicolas Bielser

September 2020

Introduction (Week 1)

Problem description / background discussion (who would be interested in this project)

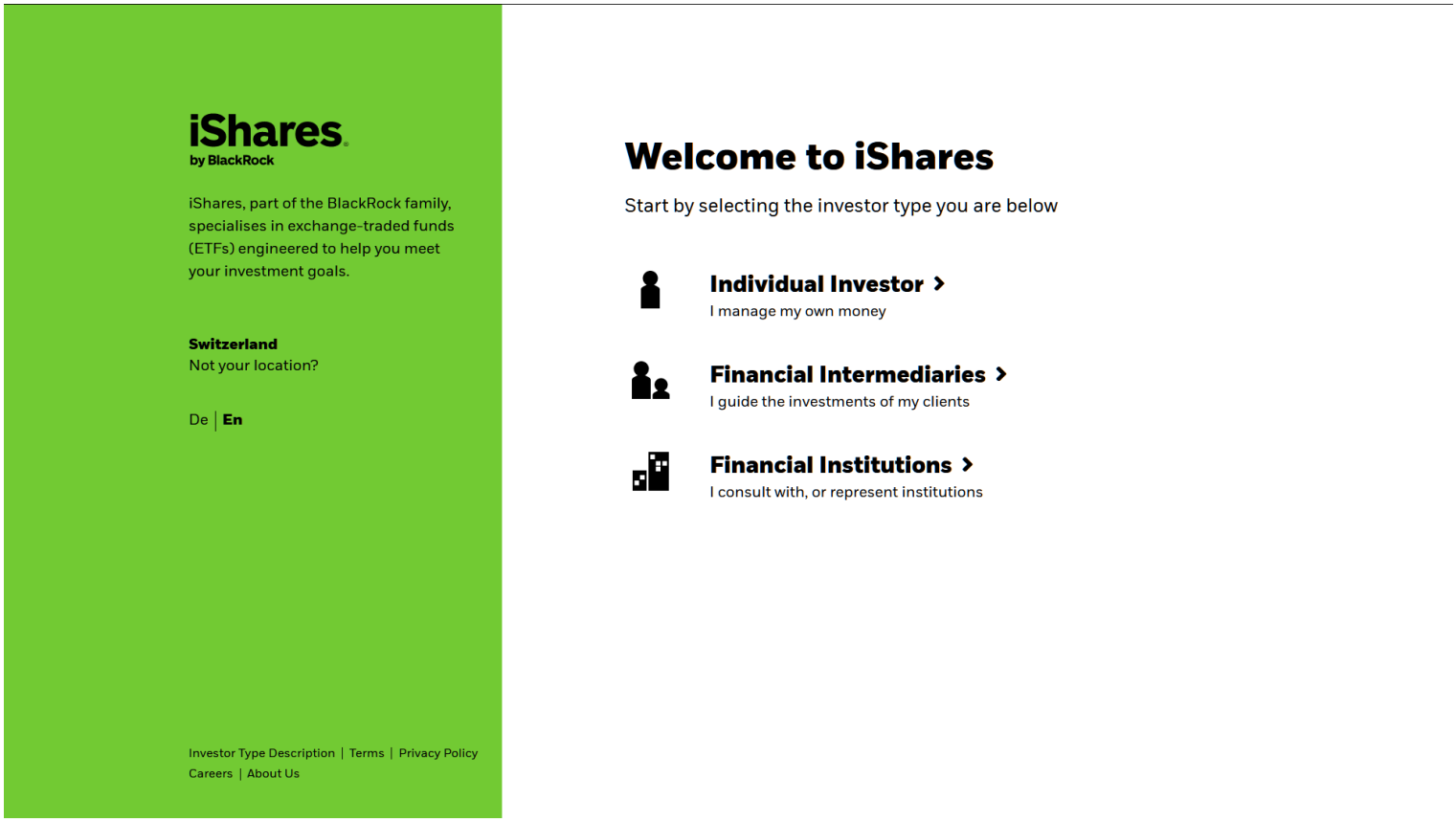
Potential investors, product management, or sales management of a bank might want to compare returns and costs of ETFs, or even exclude certain assets due to bad returns in the past (investors). As far as I am concerned, iShares (BlackRock) is one of the largest issuers for ETFs and has a product range which represents indices all over the world. The objective of this project is to visualize regions of interest and to analyze costs over the whole product range of iShares ETFs.

Data description (and use) and source (acquisition/cleaning)

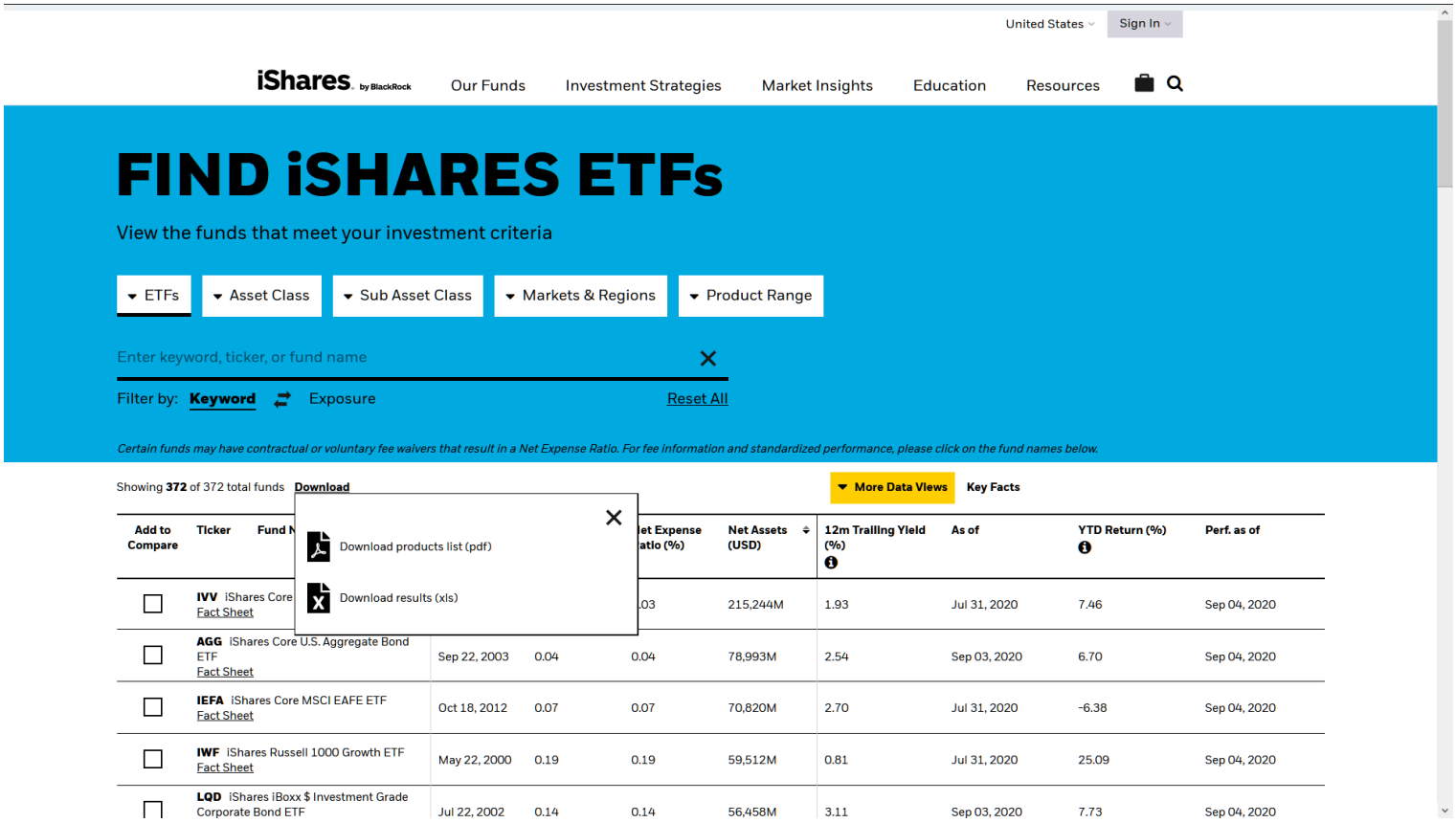
To analyze all iShares ETFs available, we **require** a data file with all products and additional information such as costs, return, and markets/indices it represents (location). I wanted to use as much data as possible to get useful information for potential investors, product management, or sales management of a bank.

As a data **source**, I first checked their homepage and found the following data set available (no registration needed; with the link below you can access the files as well if you plan to build something similar with it).

iShares ETFs USA (no specific investor type):
https://www.ishares.com/us/products/etf-investments#!type=ishares&view=keyFacts



(source iShares.com)



(source iShares.com)

Methodology (Week 2)

Exploratory data analysis and discussion/description

Choropleth Map for *YTD (%)-Return*: Location does not exist for all ETFs in the data file as such as MSCI World UCITS ETF (Ticker URTH) and they will not be represented by the choropleth map (Folium) at all. The same problem occurs with ETFs containing more than just one country (broad). An additional choropleth map might be needed to separate regions. Because we want to be more specific to countries in this project, there is no further map shown.

Cost-Return **Scatter Plot**: Just out of curiosity I wanted to see if higher priced ETFs also generate higher returns. There are plenty of time periods available (YTD, 1Y, 3Y, 5Y) in the data file, but we will only check YTD to see the current market conditions.

Box Plot for *Gross Expense Ratio (%)*: Annual GER is mentioned in the iShares Excel file. How volatility in a specific market impacts the price cannot be determined.

Monthly Return (%) **Heat Map**: Price data is not available in the iShares file. An additional source might be Google, Yahoo, or Bloomberg (if you have access to a Terminal).

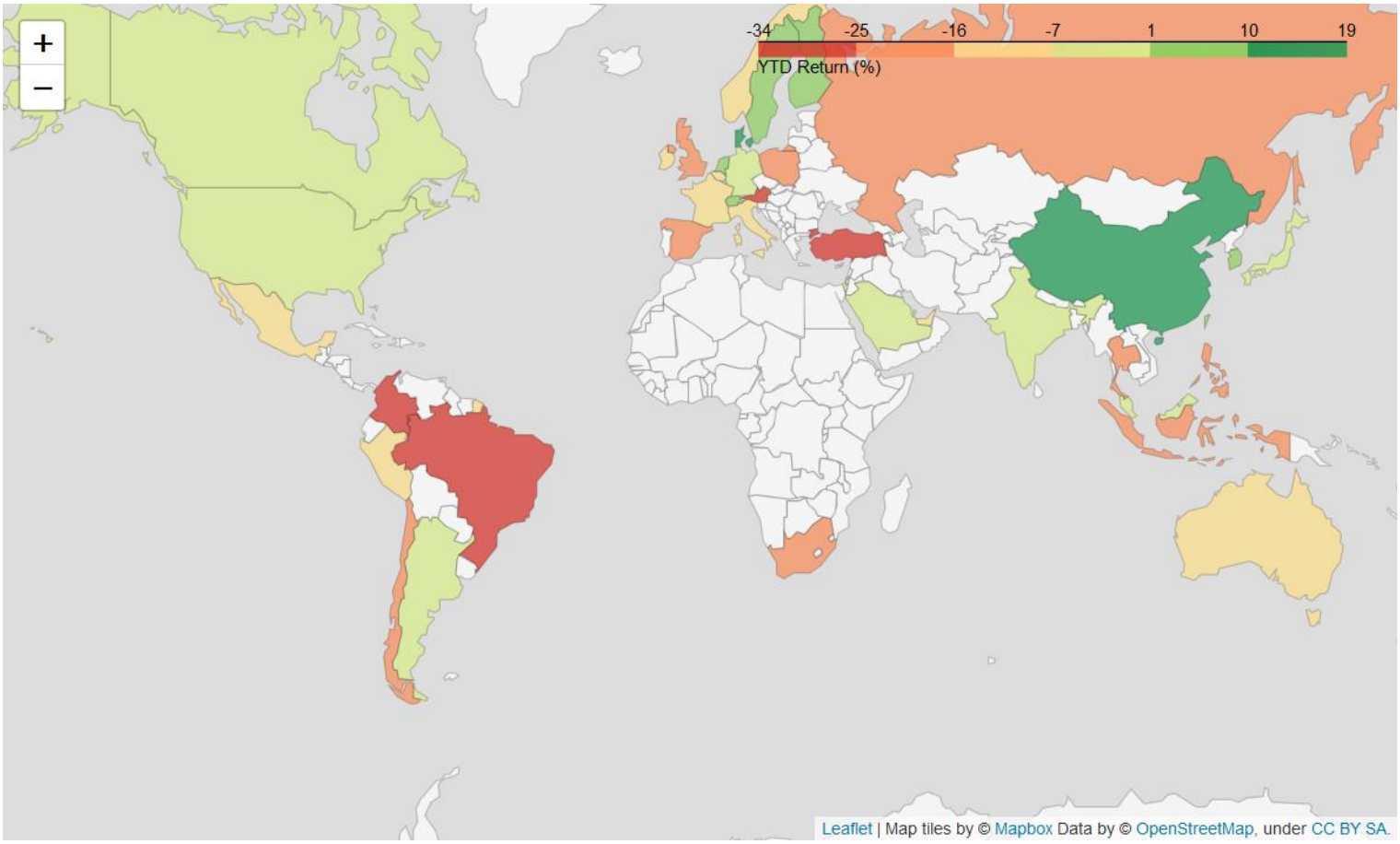
Results discussion

Choropleth Map for YTD (%)-Return: The first thing, I was thinking about, was a choropleth map to visualize regions of interest regarding positive returns. Index location and return data is available in the data set, so the choropleth map is be possible as the first tool for investors/sales management. The use for product management seems limited for YTD as chosen time period. On the longer run (e.g. 5Y) they can use it as well to see outperformers for new products/ETFs.

Return data in a choropleth map is also very supportive for the sales management department of a bank and helps identifying regions of interest for the sales of ETFs on such indices.

The choropleth map cannot visualize broad ETFs with e.g. MSCI World index because they have not a specific country/location defined in the data file. You might want to exclude regions with negative returns anyway and won't be able to see them at all in such broad ETFs. Excluding whole regions can be tricky if you have a closer look to Europe: Countries within Europe are currently (YTD as of per 04.09.2020) performing negative and positive. The same goes for Asia.

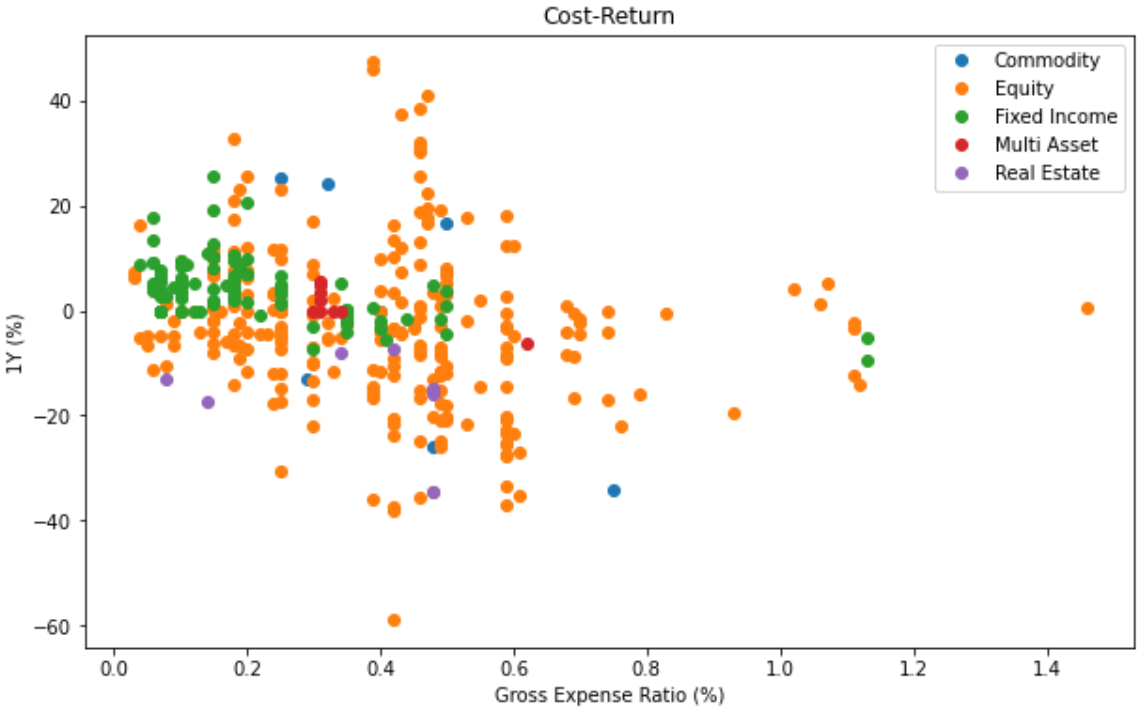
The only problem with the map is, that it represents all asset classes and not a specific such as equities, fixed income or else. If this map would be implemented in a tool, a filter by asset class is of great use.



Data as of per 04.09.2020 (data source iShares.com)

Cost-Return Scatter Plot (Matplotlib): Cost-Return comparison grouped by asset class shows how more risky assets such as equities are performing/priced compared to more traditional assets such as fixed income ETFs.

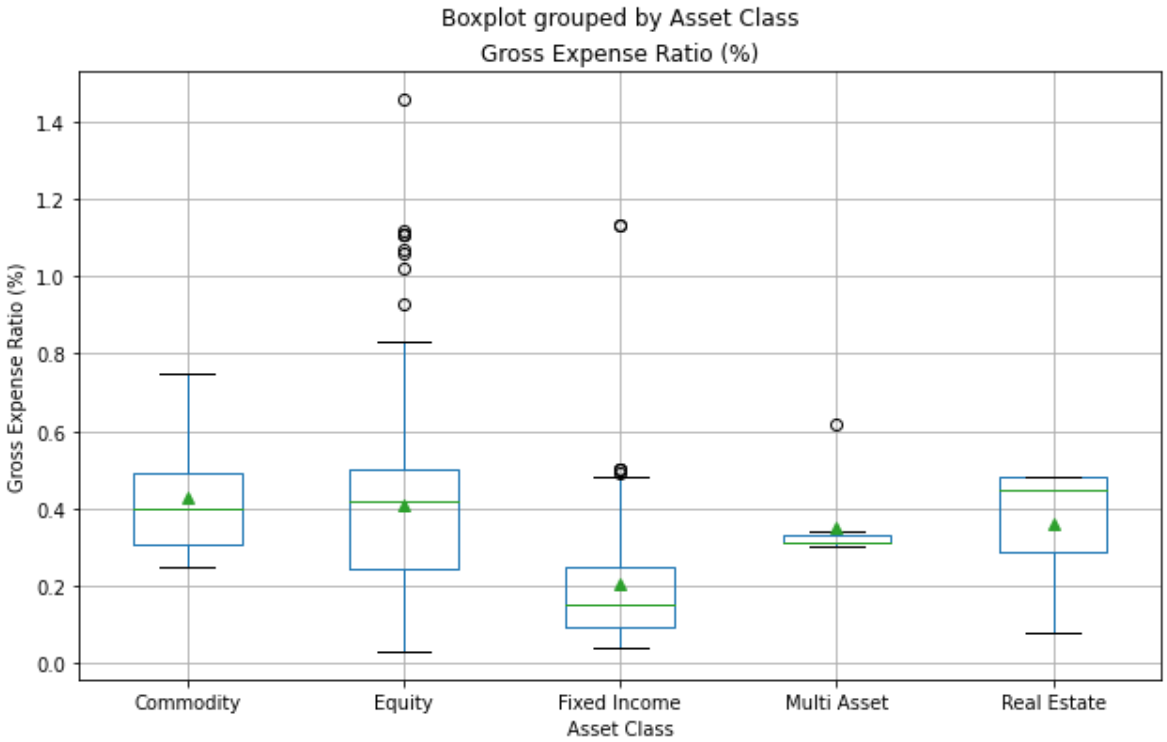
ETFs are passive investments and do not require an active fund manager which results in higher expenses. As a result, gross expense may not have a large impact to performance anyway (a few exceptions occur). For actively managed funds (e.g. hedge funds) this tool is of more use because costs have a higher impact on them (2/20 fee structure).

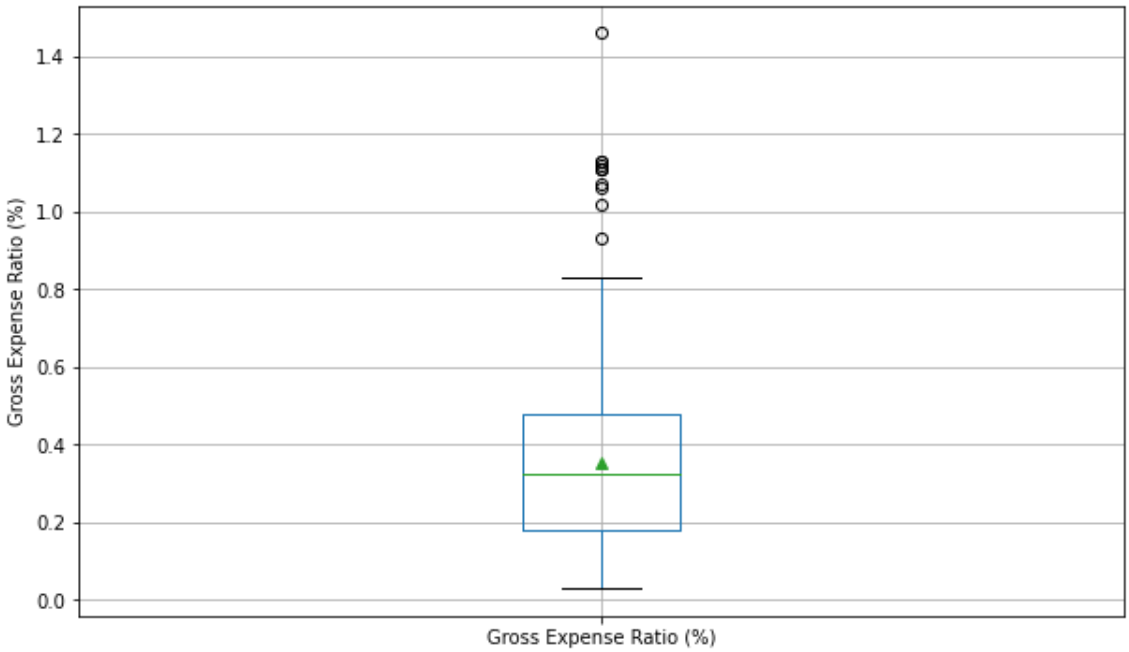


Product management might be interested in the pricing structure of iShares ETFs by asset class to price their competing ETFs accordingly or to check whether they are priced similar. On the other hand, an investor can compare his preferred/proposed ETFs as well in general – to be more specific he has to choose the ETF representing the same index out of the list.

Box Plot for Gross Expense Ratio (%): Data can be separated by asset class and used for comparison with competitor iShares or other ETFs in general and is shown in a box plot (Matplotlib) and table.

These product information if of use for product management but also for sales management to have an argument for their funds in comparison to competitors or actively managed funds.

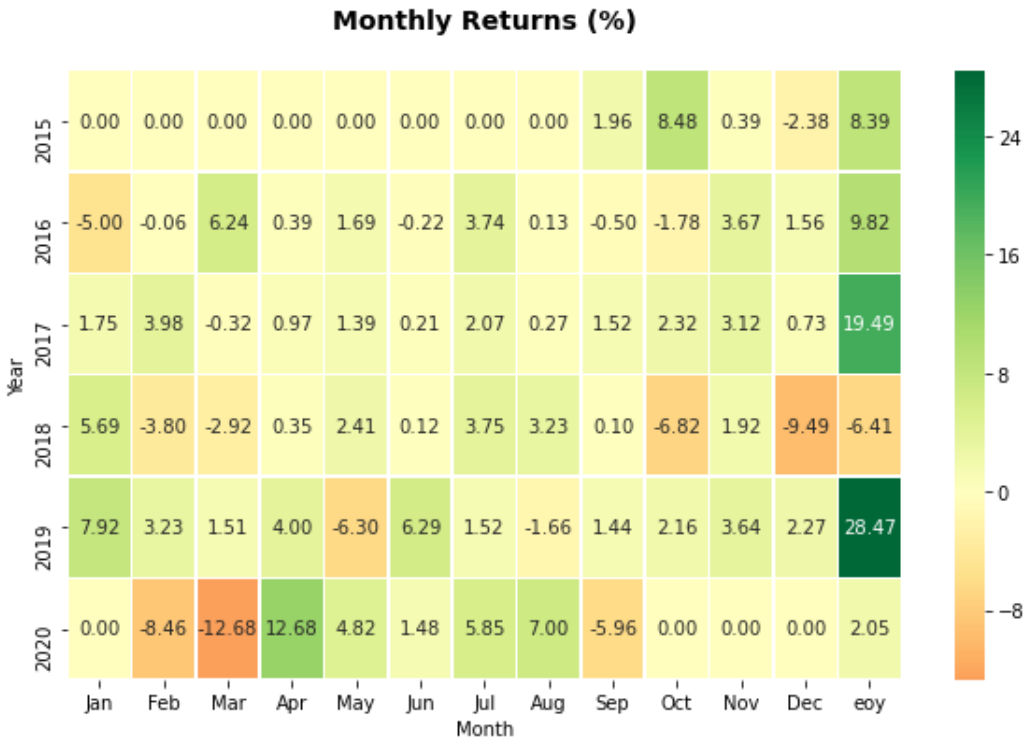




	Asset Class	Commodity	Equity	Fixed Income	Multi Asset	Real Estate	Total
Gross Expense Ratio (%)	**count**	7.000	246.000	102.000	9.000	8.000	**372.000**
	mean	0.427	0.409	0.205	0.349	0.363	**0.351**
	std	0.171	0.223	0.180	0.102	0.164	**0.226**
	min	0.250	0.030	0.040	0.300	0.080	**0.030**
	25%	0.305	0.243	0.093	0.310	0.290	**0.180**
	50%	0.400	0.420	0.150	0.310	0.450	**0.325**
	75%	0.490	0.500	0.250	0.330	0.480	**0.480**
	max	0.750	1.460	1.130	0.620	0.480	**1.460**

Return Heat Map (5Y): Performance analysis might be a useful tool for each specific ETF (use Ticker while programming) with public price data (close) from Yahoo. Consistent positive performance can be checked with a heat map for each specific ETF-Ticker. Due to the number of ETFs available in the data file (372) it is quite tricky to use for further analysis.

Blackrock's iShares largest ETF by Net Assets (USD) is iShares Core S&P 500 ETF (Ticker IVV) – due to the fact, that it represents S&P 500 it might also be one of the first choices for investors in the US market.

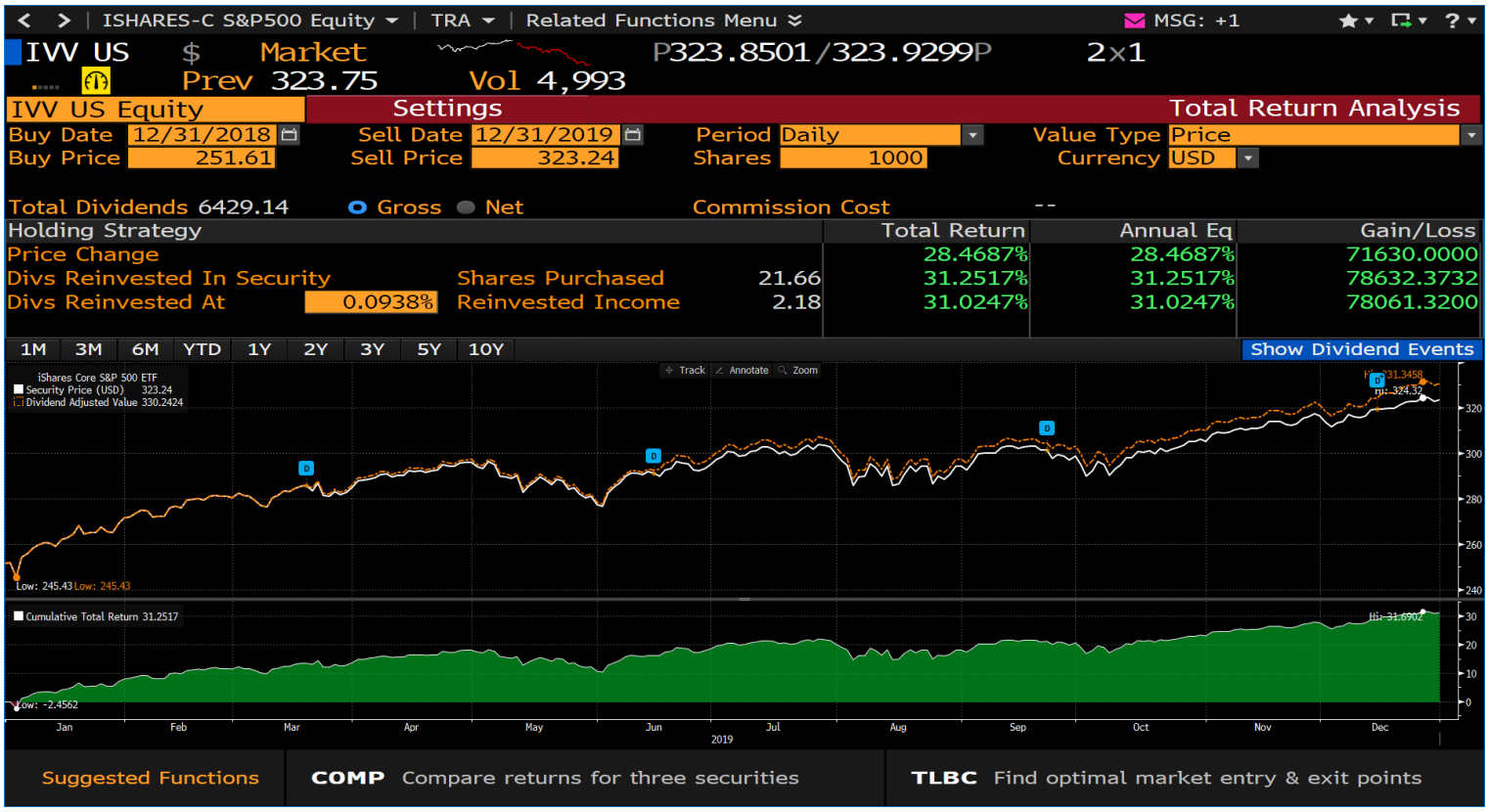



```
Date
2015-09-28    189.009995
2015-09-29    189.039993
2015-09-30    192.710007
2015-10-01    193.240005
2015-10-02    196.070007
...
2020-09-21    329.630005
2020-09-22    332.970001
2020-09-23    323.750000
2020-09-24    324.600006
2020-09-25    329.880005
Name: Close, dtype: float64
```

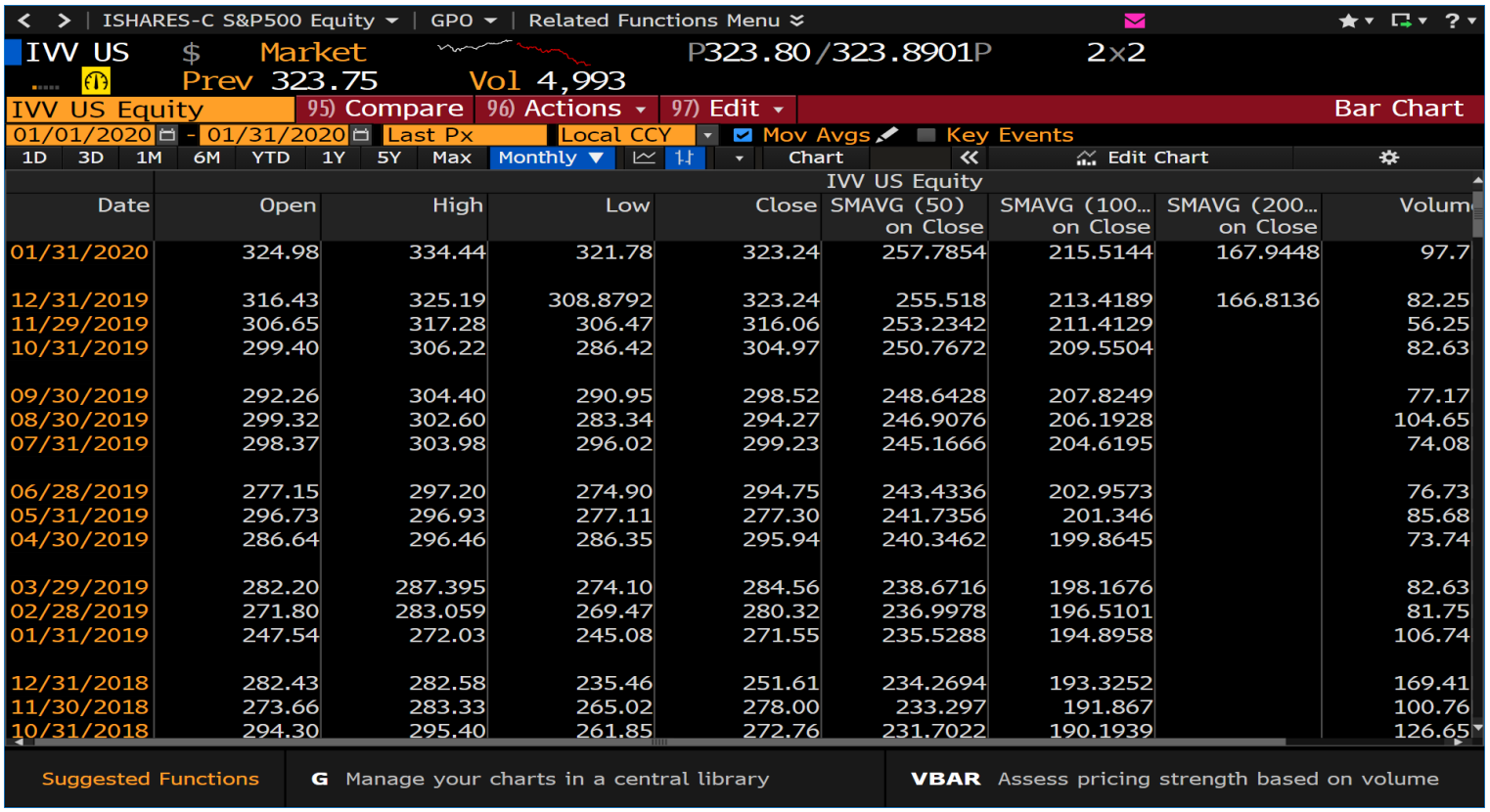
Price data feed 28.09.2015 to 25.09.2020 and represents approx. past 5 years (source Yahoo.com)

A heat map may show a way for investors how to invest in a way of just rebalancing their portfolios on an annual base.

Due to the fact, that price data (close) is from Yahoo and not from the issuer iShares itself, I wanted to verify the returns with Bloomberg:



28.47 % performance for 2019 validated (source Bloomberg Terminal)



0.00 % performance 31.12.2019 to 31.01.2020 validated (source Bloomberg Terminal)

Observation discussion and recommendations

For investors or sales management, the choropleth map gives a good indication on the markets. From a product management point of view, the annual costs of ETFs from competitor iShares can be analyzed by mean to prize own ETFs accordingly or if own ETFs are prized in/out range.

Findings of this project will help the relevant stakeholders in different ways:

- Investors or sales management of a bank can see bad performing markets on a world map and can exclude them from their portfolio.
- Product management can analyze and compare costs of one of the largest ETF issuers and can prize their own product within that range.

It's recommended to use the presented visualization methods not solely for investment decisions or product pricing, but they give a hint where to have a closer look.

Conclusion

The data file offered by iShares gave us some limitations such as missing volatility data or missing price data (close). Volatility could be added with a Bloomberg Terminal or can be calculated with more price data from Yahoo. To exclude certain regions due to negative returns in the past from your broad ETF or portfolio might not be that easy. Certain regions cannot be shorted but require more investment capital and are generating costs due to a long position in the broad ETF and an additional short position in the short ETF which does not really make sense.
On the other hand, it can be complicated to re-build a MSCI World ETF on your own without the few bad performing ETFs (some regions may not even have a liquid ETF available).

Future directions

An extension of the data file by volatility would be great for a return-risk analysis (scatter plot). With such, the report can be extended for use in portfolio management. The stakeholders might no be able to use all the presented visualizations but may want to extend them (e.g. choropleth map filtered by regions) or want to implement them on a homepage (e.g. monthly return heat map with filter for Ticker).

References

iShares ETFs: https://www.ishares.com/us/products/etf-investments#!type=ishares&view=keyFacts
Yahoo Finance (hist. price data; Ticker=IVV): https://finance.yahoo.com/quote/IVV/history?p=IVV
GitHub Repository (code, data, presentation): https://github.com/Nicolas-8/Capstone_Project_Submission-ETFs
MSCI (Index Provider): https://www.msci.com/index-solutions
2/20 Hedge Fund Mgmt. Fee Structure: https://www.investopedia.com/terms/t/two_and_twenty.asp