State Street Global Advisors UConn Senior Design Project

State Street Advisors: Joel Turcotte, Patrick Bernard

UConn Advisors: John Iacovacci, Dr. Phillip Bradford

R-Model Student Team: Kate Piotrowski, Taamir Khan,

Cloud Infrastructure Student Team: Agasti Mhatre, Souleyman Bane

Project Goal: Develop a full-scale model in R that can read various ESG data and calculate a correlation coefficient to stock price performance over time.

Project Overview:

- There are fives R files associated with the model.
    - MSCI to Price Correlation
    - GICS Sector Average
    - CQ Score to MSCI Correlation
    - CQ Score to Price Correlation
    - TruCost to Price Correlation
- There are four ESG data sets being used with this model.
    - Morgan Stanley Capital International (MSCI)
    - Global Industry Classification Standard (GICS) Sector
    - Content Quality (CQ)
    - TruCost

Project Development:

- R-Studio has been utilized as the IDE for development throughout this project.
- R-Studio Download

File Overview:

File Name	Functionality	Input(s)	Output(s)
MSCI_Price_correlation.r	• Correlates Price to MSCI Data (ESG).	1. MSCI Data (MSCIData_2014_to_2018_2021_to_2022.csv) 2. Price Data (PriceData_2012_2022.csv)	1. Table of Correlations (10_Years_Correlation_Table.csv) 2. MSCI and Price Data Merged (MSCI_Price_Data.csv)
GICS_Sector_Avg.r	• Calculates the GICS average per sector. • Determines outliers. • Correlates GICS Sector Average to MSCI Data.	1. Price Data (Price_Data3_ISIN.csv) 2. MSCI Correlations (fixed_10_Years_Correlation_Table.csv) 3. GICS Sectors (GICS_Sectors.csv)	1. GICS Sector Averages (GICS_Sector_MSCI_Correlation_Avg.cs) 2. Outliers by GICS Sector Avg (Outliers_By_GICS_Sector.csv)
CQ_Score_to_ESG_MSCI.r	• Correlates CQ score to MSCI Data (ESG).	1. CQ Scores Data (cq.csv) 2. Cleaned CQ Data (Removed_CQ_Data.csv) 3. MSCI Data (ESG_Price_Date_Data.csv)	1. CQ Score to MSCI Correlation (CQ_Score_to_ESG_MSCI_Correlation.csv)
CQ_Score_to_Price.r	• Correlates CQ score to Price.	1. CQ Scores Data (cq.csv) 2. Cleaned CQ Data (Removed_CQ_Data.csv) - This includes price data	1. CQ Score to Price Correlation (CQ_Score_to_Price_Correlations.csv)
truCost_correlation.r	• Correlates Price to TruCost.	1. TruCost Data (truCost.csv) 2. Price Data (PriceData_2012_2022.csv)	1. TruCost to Price Correlations (TruCost_to_Price_Correlation.csv)

MSCI to Price Correlation Model

Part 1 Input:

stock_price_data from PriceData_2012_2022.csv

DATE_DIM_ID	BB_TICKER_CD	ISIN_CD	PBD_PRICE_AMT	year_month
2014-08-29	ZTS US	US98978V1035	35.44	2014-08
2014-09-04	ZTS US	US98978V1035	35.51	2014-09
2014-09-10	ZTS US	US98978V1035	36.24	2014-09
2014-09-12	ZTS US	US98978V1035	36.18	2014-09

Part 2

Find the average monthly price per security

ISIN_CD	year_month	avg_price
AN8068571086	2012-11	69.69950
AN8068571086	2012-12	70.65000
AN8068571086	2013-01	75.06095
AN8068571086	2013-02	78.60211

Part 3 Input:

msci_data from MSCIData_2014_to_2018_2021_to_2022.csv

ASOF_DATE	ISSUER_ISIN	INDUSTRY_ADJUSTED_SCORE	WEIGHTED_AVERAGE_SCORE	IVA_INDUSTRY	ENVIRONMENTAL_PILLAR_SCORE	GOVERNANCE_PILLAR_SCORE	SOCIAL_PILLAR_SCORE	year_month
2014-01-01	US0236081024	2.50	3.9	Utilities	3.2	8.5	6.9	2014-01
2014-01-01	US09247X1019	8.00	6.5	Asset Management	5.0	3.0	6.8	2014-01
2014-01-01	US7433151039	4.80	4.5	Property & Casualty Insurance	5.2	6.6	3.1	2014-01
2014-01-01	US1101221083	5.10	4.2	Pharmaceuticals	4.8	3.8	4.4	2014-01
2014-01-01	US2479162081	7.30	5.4	Oil & Gas Exploration & Production	5.0	7.1	4.3	2014-01

Part 4:

Combine average price and ESG score

ISIN_CD	year_month	avg_price	avg_environmental_score	avg_social_score	avg_governance_score
AN8068571086	2014-12	84.97500	6.4	5.3	5.7
AN8068571086	2015-01	81.43250	6.4	5.3	5.7
AN8068571086	2015-02	85.89789	6.4	5.3	5.0
AN8068571086	2015-03	82.61773	6.4	5.3	5.0

Part 5 Output:

Calculated correlations per ESG score to price (correlations_msci.csv)

ISIN_CD	environmental_score_correlation	social_score_correlation	governance_score_correlation
AN8068571086	-0.136991869	-0.46768064	-0.442566207
BMG3223R1088	-0.777577153	0.76521218	-0.518748805

GICS Sector Average Model

Part 1 Input:

Price_Data from PriceData_2012_2022.csv

DATE_DIM_ID	BB_TICKER_CD	ISIN_CD	PBD_PRICE_AMT	year_month
2014-08-29	ZTS US	US98978V1035	35.44	2014-08
2014-09-04	ZTS US	US98978V1035	35.51	2014-09
2014-09-10	ZTS US	US98978V1035	36.24	2014-09
2014-09-12	ZTS US	US98978V1035	36.18	2014-09

Part 2 Input:

MSCI_Correlation_Data from fixed_10_Years_Correlation_Table.csv

isinID	governancePillarScore	environmentalPillarScore	socialPillarScore
US0605051046	0.9213322	0.61607059	0.240115647
US4410601003	0.9083940	-0.79047968	-0.334921731
US37247D1063	0.8986127	0.90223321	-0.890752266
US46625H1005	0.8972508	0.49147501	0.041393895

Part 3 Input:

gics_sector_df from GICS_Sectors.csv

isin_ID	sector_level_1
US8760301072	Consumer Discretionary
US00971T1016	Information Technology
US0010551028	Financials
US42250P1030	Real Estate

Part 4 Output:

joined together MSCI to Price correlation with GICS Sector

isin_ID	environmental_correlation	social_correlation	goverance_correlation	sector_level_1
AN8068571086	-0.810610746	-0.68733490	-0.680363829	Energy
BMG3223R1088	-0.024003061	0.06896253	-0.567242870	Financials
BMG491BT1088	-0.585447206	0.04789883	0.011421061	Financials
BMG667211046	-0.749440974	-0.71909902	-0.916910533	Consumer Discretionary
CH0044328745	-0.853012567	0.39099629	-0.404151996	Financials

Part 5 Output:

entire_joined_sectors_avg (calculates the avg correlation by GICS Sector)

sector_level_1	environmental_correlation	social_correlation	goverance_correlation
Communication Services	0.30697120	0.2397773	-0.30882609
Consumer Discretionary	-0.10906882	0.2202777	-0.04968211
Consumer Staples	0.21456964	0.2618437	-0.01998541
Energy	0.22552648	-0.0272824	0.05848824
Financials	0.09299335	0.1308040	-0.03364343
Health Care	0.13422930	0.1944427	0.02258770
Industrials	0.07753924	0.2086654	-0.03177065
Information Technology	-0.02986816	0.1659122	-0.16472368
Materials	0.11714548	-0.1056803	0.03065386
Real Estate	0.27406580	-0.1492297	-0.31957017
Utilities	0.50321401	0.2702451	-0.13130124

CQ Score to MSCI (ESG) Correlations Model

Part 1 Input:

df from cq.csv

date	ticker	isin_ID	price	cq_score
20160114	ABT US	US0028241000	41.1	8.6
20160104	ABT US	US0028241000	42.93	8.6
20160223	ABT US	US0028241000	38.39	8.6
20160209	ABT US	US0028241000	37	8.6

Part 2:

Separate Month and Year for CQ Score (cq_agg)

ticker	Month	Year	Month_Year_Together	price	cq_score
AA US	01	2015	01/2015	15.662500	6.760000
AA US	01	2016	01/2016	7.532105	5.636842
AA US	02	2015	02/2015	15.829474	6.900000
AA US	02	2016	02/2016	8.095000	5.500000
AA US	03	2015	03/2015	13.537273	6.900000

Part 3 Input:

Read MSCI data from ESG_Price_Date_Data.csv

ticker	isin_ID	price	month	year	price.1	environmental_pillar_score	social_pillar_score	governance_pillar_score
ZTS US	US98978V1035	35.17167	8	2014	35.17167	4.8	7.7	3.0
ZTS US	US98978V1035	36.33333	9	2014	36.33333	4.8	7.7	3.0

Part 4:

Combine MSCI Score and CQ Score data frames to create df4

ticker.x	Month	Year	Month_Year_Together	cq_score	isin_ID	price.1	environmental_pillar_score	social_pillar_score	governance_pillar_score
AA US	01	2015	01/2015	6.760000	US0138171014	15.662500	4.0	6.6	5.1
AA US	02	2016	02/2016	5.500000	US0138171014	8.095000	4.8	5.8	5.0
AA US	05	2015	05/2015	6.600000	US0138171014	13.372000	4.6	7.0	5.3
AA US	05	2016	05/2016	5.500000	US0138171014	9.578571	4.8	5.8	5.5
AA US	07	2016	07/2016	5.500000	US0138171014	10.343000	4.8	5.8	5.1
AA US	08	2016	08/2016	5.500000	US0138171014	10.310000	4.8	5.8	4.9
AA US	12	2015	12/2015	5.700000	US0138171014	9.380000	4.8	5.8	4.3

Part 5 Output:

Correlate each MSCI Score to CQ Score (CQ_Score_to_ESG_MSCI_Correlation)

ticker.x	governancePillarScore	environmentalPillarScore	socialPillarScore
AA US	0.19055302	-0.853243143	0.942696582
AAP US	-0.64357626	-0.477418234	-0.394423797
AAPL US	-0.46356127	-0.530617373	0.163287973
ABBV US	-0.00473079	-0.441724451	-0.509702357

CQ Score to Price Correlations Model

The model follows the same steps as the CQ Score to MSCI (ESG) Correlations Model outlined above. The difference is the output as the correlation coefficient is CQ to Price versus MSCI.

Final Output:

Correlate each Price to CQ Score (correlations)

ticker	correlation_coefficient
AA US	0.427763256
AAL US	0.157085577
AAP US	0.114233007
AAPL US	0.107405482
ABBV US	-0.587949779
ABMD US	0.036775511

TruCost to Price Correlations Model

Part 1 Input:

stock_price_data from PriceData_2012_2022.csv

DATE_DIM_ID	BB_TICKER_CD	ISIN_CD	PBD_PRICE_AMT	year_month
2014-08-29	ZTS US	US98978V1035	35.44	2014-08
2014-09-04	ZTS US	US98978V1035	35.51	2014-09
2014-09-10	ZTS US	US98978V1035	36.24	2014-09

Part 2:

Calculate monthly average TruCost and Price (monthly_avg_tru_cost) & (monthly_avg_price)

ISIN_CD	year_month	avg_price
AN8068571086	2012-11	69.69950
AN8068571086	2012-12	70.65000
AN8068571086	2013-01	75.06095

ISIN	GICSSECTORNAME	year_month	avg_scope1	avg_scope2	avg_first_tier
AN8068571086	Energy	2017-04	3.962113e+01	16.3495419	190.182072
AN8068571086	Energy	2020-10	4.582159e+01	18.6291370	199.882287
AN8068571086	Energy	2022-08	4.099200e+01	25.1830000	201.182000

Part 3:

Combine and clean TruCost data (combined_data)

ISIN_CD	year_month	avg_price	avg_scope1	avg_scope2	avg_first_tier
AN8068571086	2017-04	76.78789	39.6211296	16.349542	190.182072
AN8068571086	2017-05	71.28864	39.6211296	16.349542	190.182072
AN8068571086	2017-06	67.57864	39.6211296	16.349542	190.182072

Part 4:

Calculate correlation coefficient between Price and TruCost (correlations_with_ticker)

ISIN_CD	BB_TICKER_CD	carbon_intensity_direct_and_first_tier_indirect_correlation	carbon_emissions_scope_1_correlation	carbon_emissions_scope_2_correlation
US98978V1035	ZTS US	0.664947898	-0.877261635	0.774581142
US98956P1021	ZMH US	0.283502686	0.175567421	0.421709955
US9897011071	ZION US	-0.219434326	-0.208181034	-0.186459612
US9892071054	ZBRA US	0.139517132	0.139517132	0.139517132
US98956P1021	ZBH US	0.283502686	0.175567421	0.421709955