



# Sector Dominance: Group Project 1

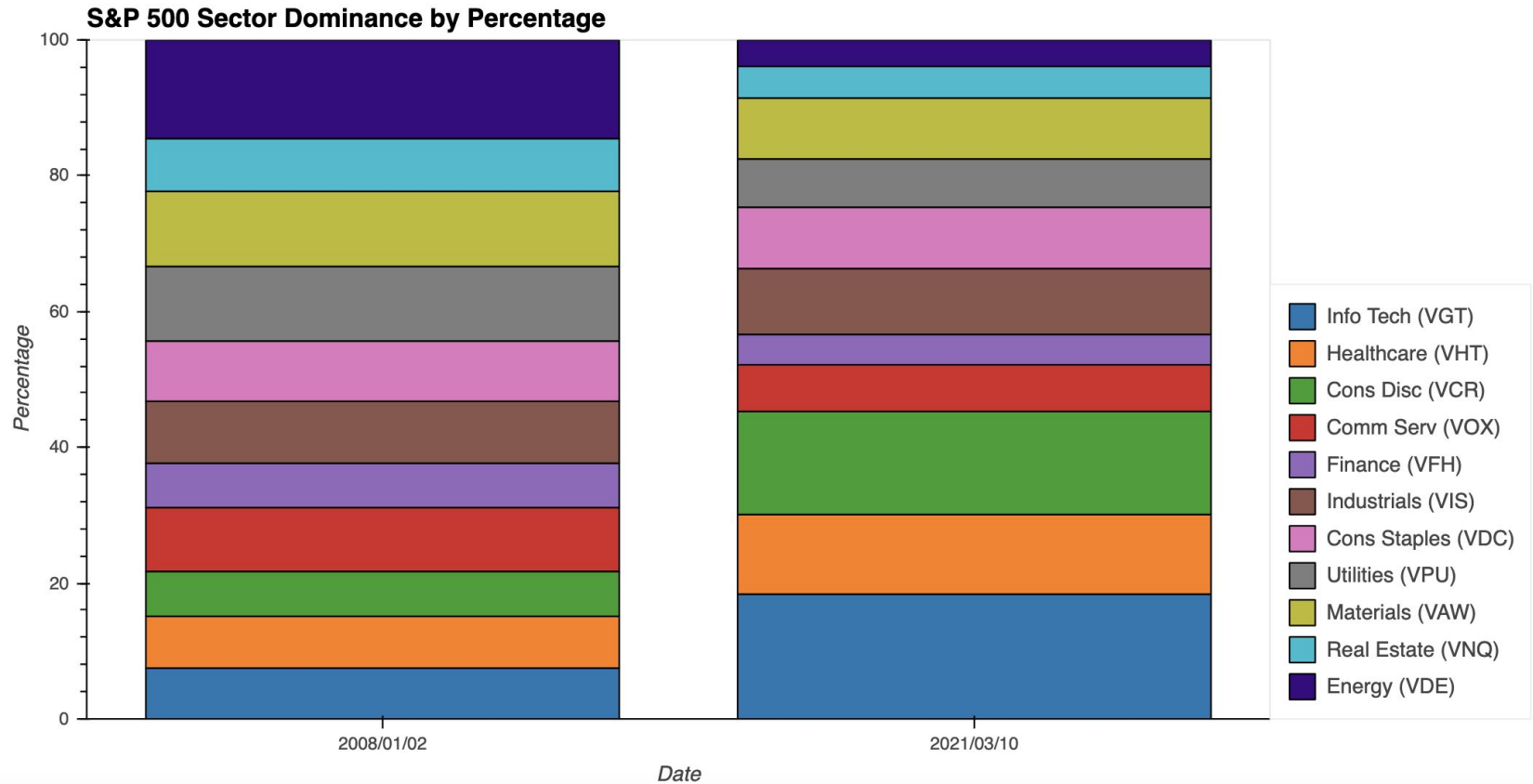
Matt Epler  
Yoko Yamamoto  
Scott Alexander  
Drew Marnell



Original hypothesis:

The technology sector has dominated the economy in the past  
13 years and it will continue to do so

# Overall Sector Growth



# Sector ETFs



U.S. Sector Fund Name	Ticker
Vanguard Consumer Discretionary ETF	VCR
Vanguard Consumer Staples ETF	VDC
Vanguard Energy ETF	VDE
Vanguard Financials ETF	VFH
Vanguard Health Care ETF	VHT
Vanguard Industrials ETF	VIS
Vanguard Information Technology ETF	VGIT
Vanguard Materials ETF	VAW
Vanguard REIT Index ETF	VNQ
Vanguard Telecommunications Services ETF	VOX
Vanguard Utilities ETF	VPU

```

# Set start and end dates
start_day = '2008-01-02'
final_day = '2021-03-10'

# Set tickers
ticker = ["VGT", "VHT", "VCR", "VOX", "VFH", "VIS", "VDC", "VPU", "VAW", "VNQ", "VDE", "SPY"]

# Set timeframe to '1D' for Alpaca API
timeframe = "1D"

# Get current closing prices for sectors
ticker_data = []

# Loop 1000 days at a time until final day is reached
while str(start_day) < str(final_day):
    end_day = np.busday_offset(start_day, 1000, roll='forward')
    if str(end_day) > str(final_day):
        end_day = final_day
    # Pull in data from Alpaca API using start and end dates
    data = alpaca_api.get_barset(
        ticker,
        timeframe,
        start = pd.Timestamp(start_day, tz="America/New_York").isoformat(),
        end = pd.Timestamp(end_day, tz="America/New_York").isoformat(),
        limit = 1000
    ).df

    ticker_data.append(data)

    start_day = np.busday_offset(end_day, 1, roll='forward')

sectors_df = pd.concat(ticker_data, axis="rows", join="inner")
sectors_df

```

START: 2008-01-02	END: 2011-11-02
START: 2011-11-03	END: 2015-09-03
START: 2015-09-04	END: 2019-07-05
START: 2019-07-08	END: 2021-03-10

time	SPY					VAW					...	VOX				
	open	high	low	close	volume	open	high	low	close	volume	...	open	high	low	close	volume
2008-01-02 00:00:00-05:00	146.53	146.99	143.8800	144.5500	175801806	87.61	88.26	86.68	87.18	43027.0	...	74.99	75.7000	73.820	74.140	28650.0
2008-01-03 00:00:00-05:00	144.92	145.49	144.1200	144.4600	111357758	87.80	88.64	87.58	87.97	18500.0	...	74.10	74.5400	73.670	73.860	24900.0
2008-01-04 00:00:00-05:00	143.34	143.44	140.9200	140.9300	195945264	86.46	86.63	85.34	85.43	50000.0	...	73.54	73.5500	72.130	72.160	29400.0
2008-01-07 00:00:00-05:00	141.81	142.23	140.1000	141.4600	220398905	85.66	86.00	83.22	84.35	155001.0	...	72.44	73.3600	72.070	73.000	94500.0
2008-01-08 00:00:00-05:00	142.10	142.90	138.6300	138.8100	303249870	85.12	85.71	82.64	82.64	104400.0	...	73.50	74.4700	69.810	70.490	110000.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2021-03-04 00:00:00-05:00	381.22	384.00	371.8800	376.7300	168203161	163.70	164.46	157.89	160.65	163250.0	...	128.38	130.3913	126.100	127.850	147853.0
2021-03-05 00:00:00-	380.46	384.76	372.6400	383.5950	133661539	162.42	165.13	158.87	164.62	126019.0	...	128.98	130.8660	126.465	130.510	246036.0

*# Create and empty DataFrame for closing prices*

```
sectors_closing = pd.DataFrame()
```

*# Fetch the closing prices of all sectors*

```
sectors_closing["Info Tech (VGT)"] = sectors_df["VGT"]["close"]
sectors_closing["Healthcare (VHT)"] = sectors_df["VHT"]["close"]
sectors_closing["Cons Disc (VCR)"] = sectors_df["VCR"]["close"]
sectors_closing["Comm Serv (VOX)"] = sectors_df["VOX"]["close"]
sectors_closing["Finance (VFH)"] = sectors_df["VFH"]["close"]
sectors_closing["Industrials (VIS)"] = sectors_df["VIS"]["close"]
sectors_closing["Cons Staples (VDC)"] = sectors_df["VDC"]["close"]
sectors_closing["Utilities (VPU)"] = sectors_df["VPU"]["close"]
sectors_closing["Materials (VAW)"] = sectors_df["VAW"]["close"]
sectors_closing["Real Estate (VNQ)"] = sectors_df["VNQ"]["close"]
sectors_closing["Energy (VDE)"] = sectors_df["VDE"]["close"]
sectors_closing["S&P 500 (SPY)"] = sectors_df["SPY"]["close"]
```

*# Clean up data*

```
sectors_closing = sectors_closing.sort_index()
sectors_closing = sectors_closing.dropna()
sectors_closing = sectors_closing.drop_duplicates()
sectors_closing = sectors_closing.rename_axis("Date")
```

*# Copying data to csv for reference*

```
sectors_closing.to_csv("sector_closing_data.csv")
```

*# Display sample data*

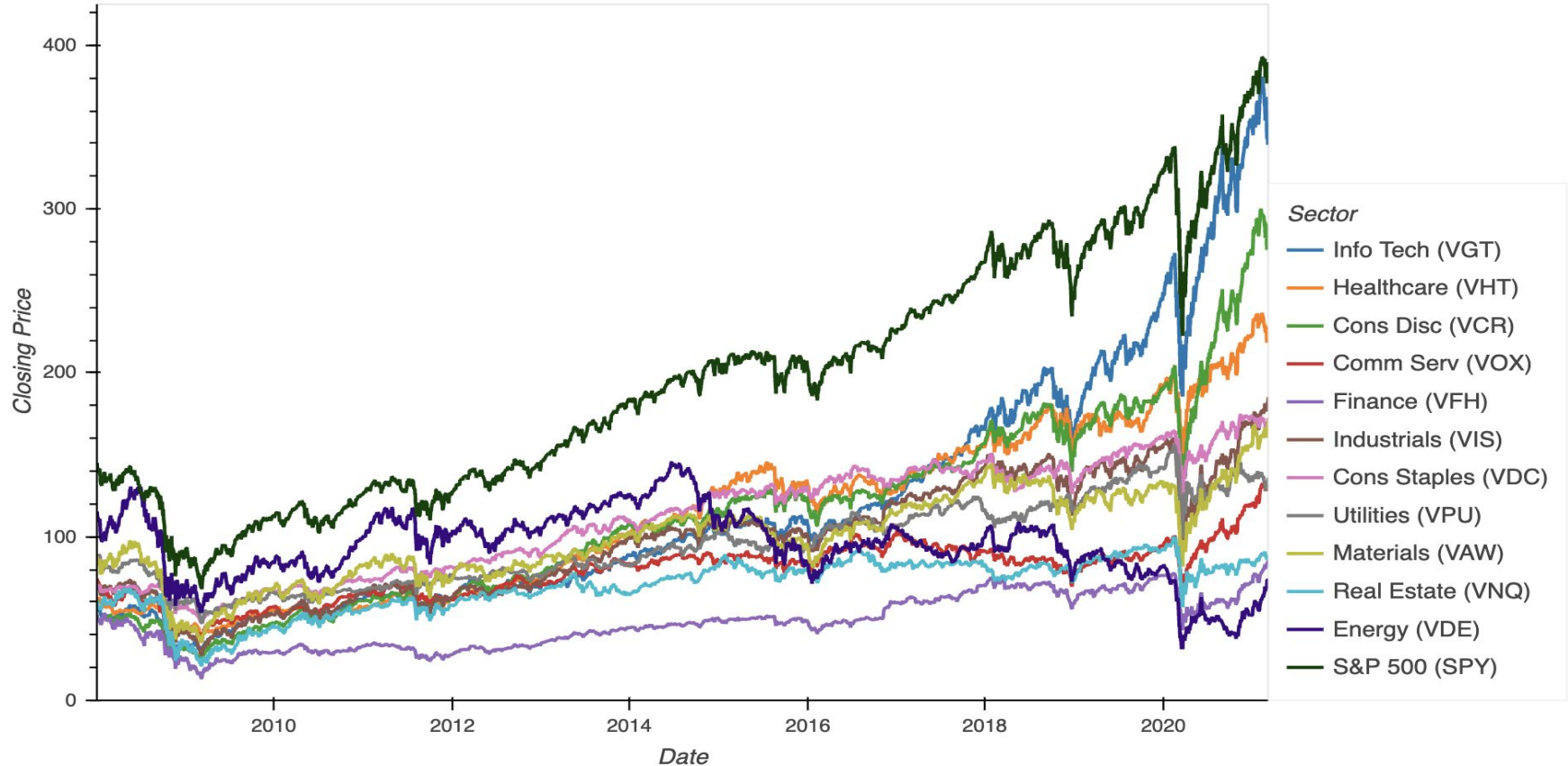
```
sectors_closing
```



	Info Tech (VGT)	Healthcare (VHT)	Cons Disc (VCR)	Comm Serv (VOX)	Finance (VFH)	Industrials (VIS)	Cons Staples (VDC)	Utilities (VPU)	Materials (VAW)	Real Estate (VNQ)	Energy (VDE)	S&P 500 (SPY)
Date												
2008-01-02 00:00:00-05:00	59.01	60.14	52.12	74.140	51.48	71.95	69.990	86.60	87.18	61.27	114.530	144.5500
2008-01-03 00:00:00-05:00	58.71	60.40	51.33	73.860	51.03	72.12	69.620	86.62	87.97	59.06	115.260	144.4600
2008-01-04 00:00:00-05:00	56.23	59.76	49.73	72.160	49.62	70.20	69.010	86.67	85.43	56.94	112.360	140.9300
2008-01-07 00:00:00-05:00	55.70	60.87	49.87	73.000	49.85	69.66	70.200	88.35	84.35	57.52	110.890	141.4600
2008-01-08 00:00:00-05:00	54.19	61.25	48.82	70.490	48.05	67.80	69.830	88.26	82.64	55.49	109.340	138.8100
...	...	...	...	...	...	...	...	...	...	...	...	...
2021-03-04 00:00:00-05:00	343.04	218.27	275.10	127.850	81.95	176.57	165.350	128.95	160.65	85.92	70.790	376.7300
2021-03-05 00:00:00-05:00	348.40	222.33	278.08	130.510	83.50	180.73	169.381	131.36	164.62	86.90	73.499	383.5950
2021-03-08 00:00:00-05:00	339.23	220.97	279.05	128.860	84.70	182.28	170.000	133.33	167.07	87.61	73.520	381.7400
2021-03-09 00:00:00-05:00	352.26	223.65	287.96	130.500	84.05	182.25	170.250	134.82	168.18	88.19	72.070	387.1100
2021-03-10 00:00:00-05:00	350.83	223.91	289.78	131.305	85.61	185.20	172.430	135.86	171.19	89.12	74.220	389.5199
3314 rows x 12 columns												

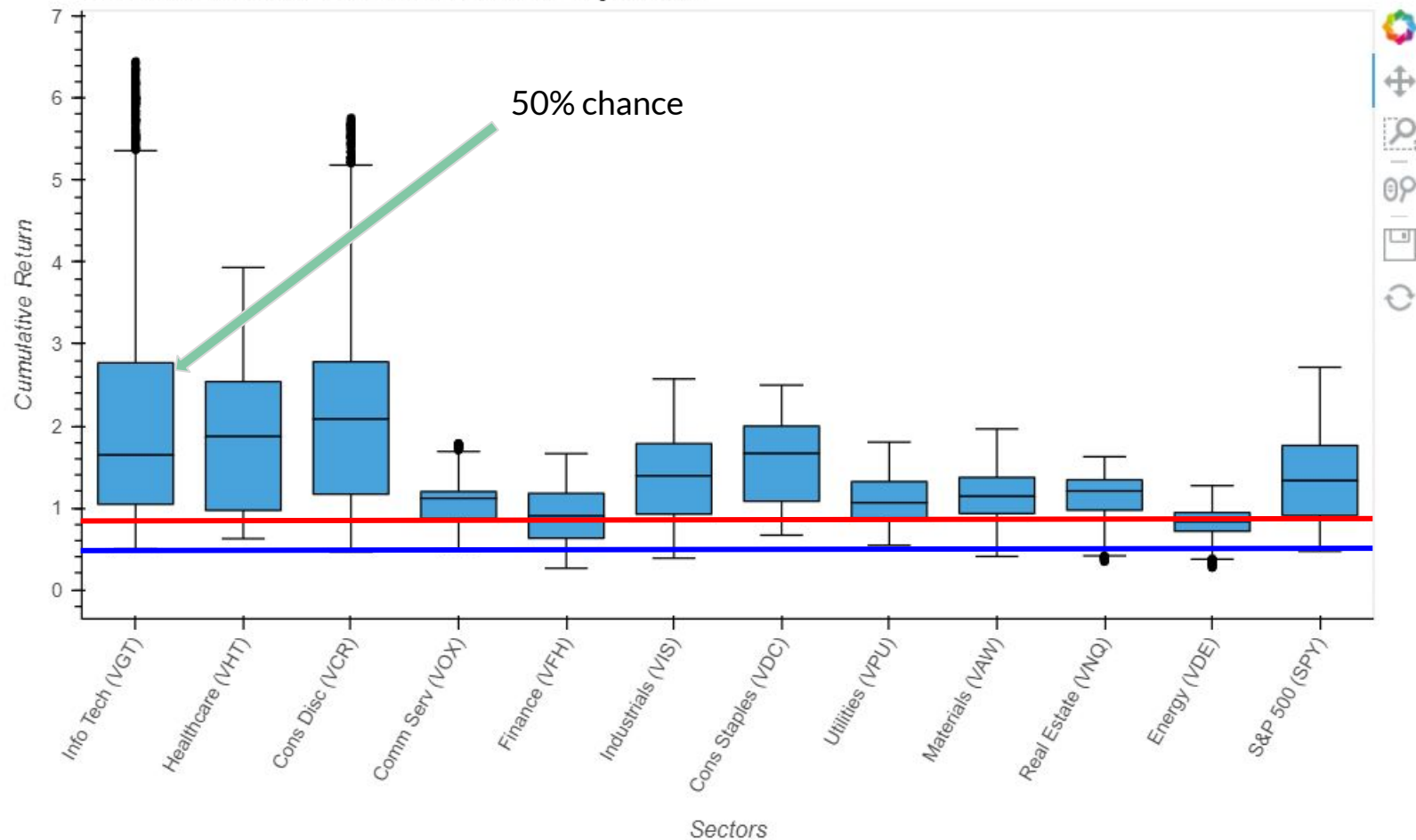


**S&P 500 Closing Prices by Sector From 2008-2021**

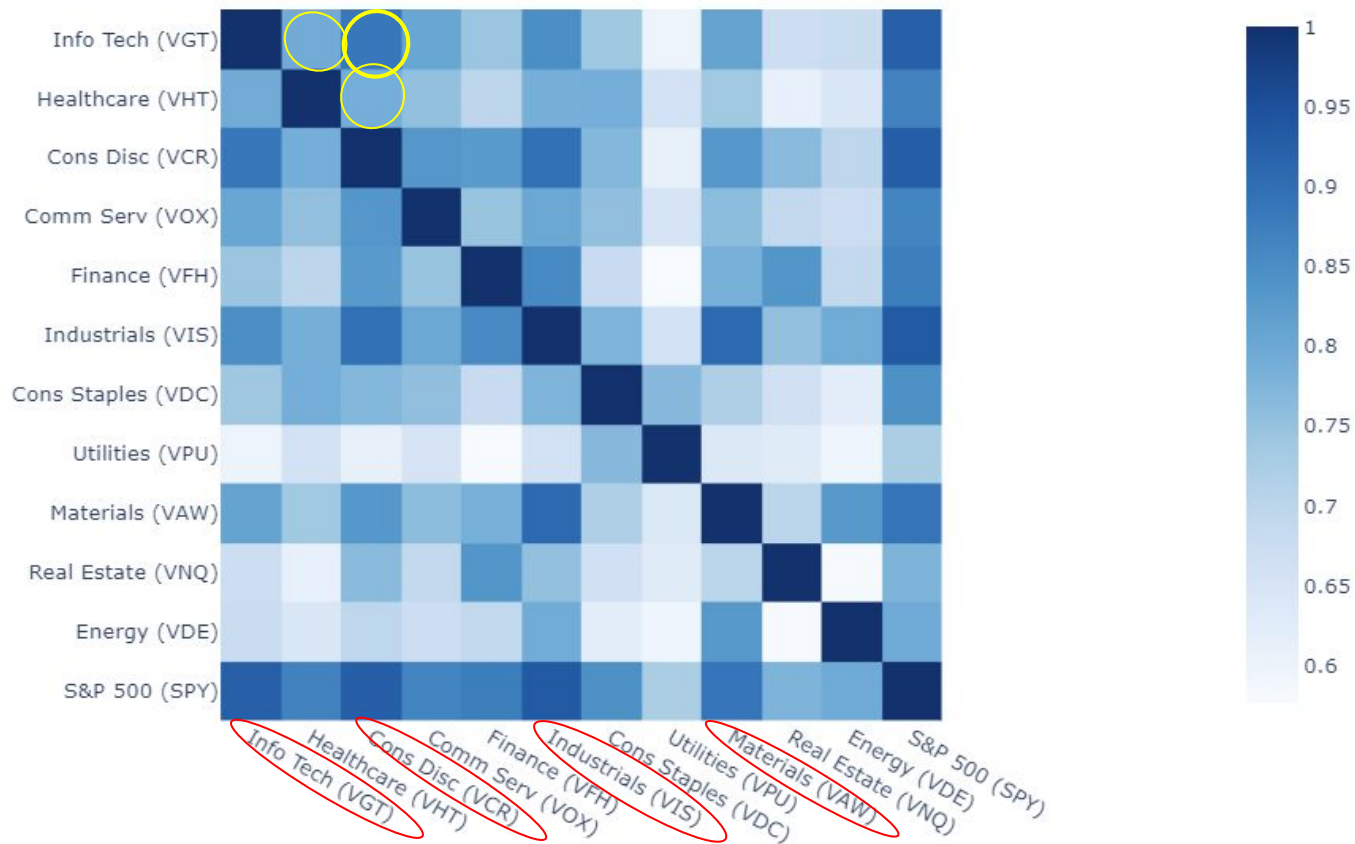


- Energy, Materials & Utilities have highest closing prices in 2008
- Info Tech, Consumer Discretionary & Healthcare have highest closing prices in 2021

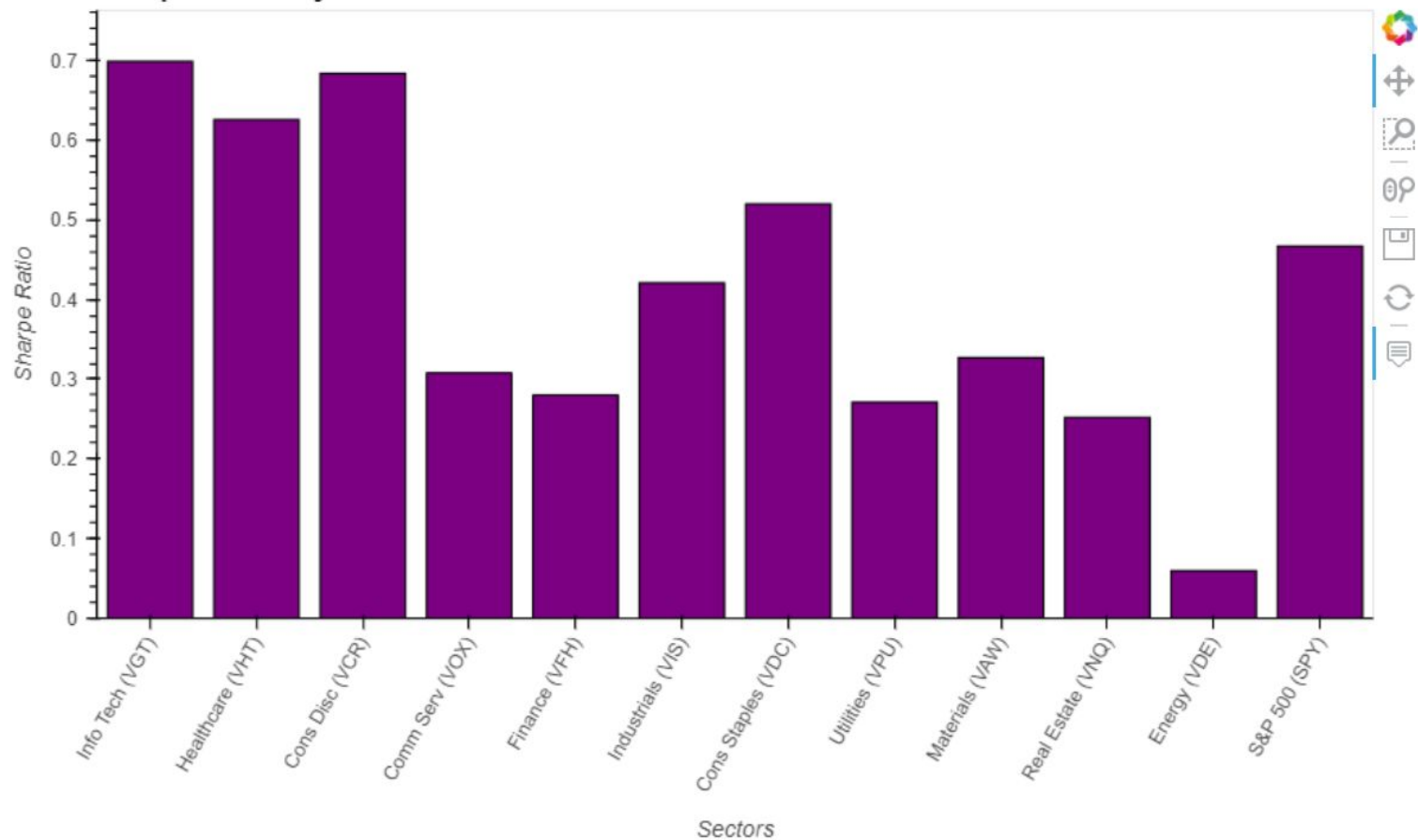
# Cumulative Returns From 2008-2021 by Sector



## All Sectors Correlation Heatmap

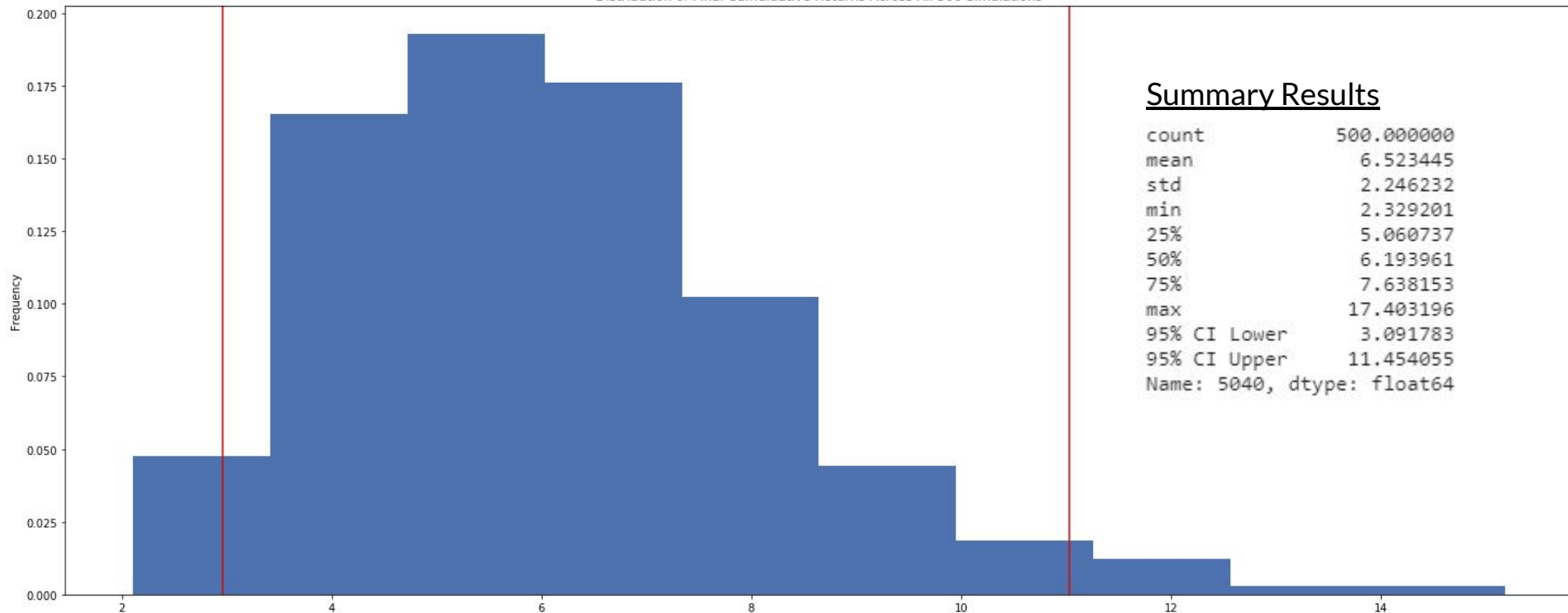


### Sharpe Ratios by Sector



# Unweighted Portfolio

Distribution of Final Cumulative Returns Across All 500 Simulations

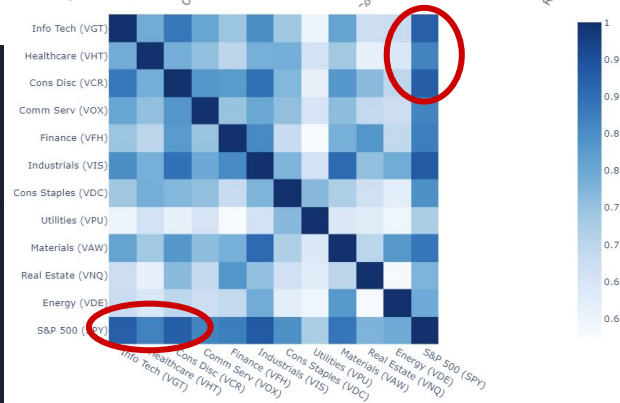


# Creating our Custom Portfolio



Highest Risk/Reward ratio:

- ❖ Information Technology (VGT)
- ❖ Healthcare (VHT)
- ❖ Consumer Discretionary (VCR)

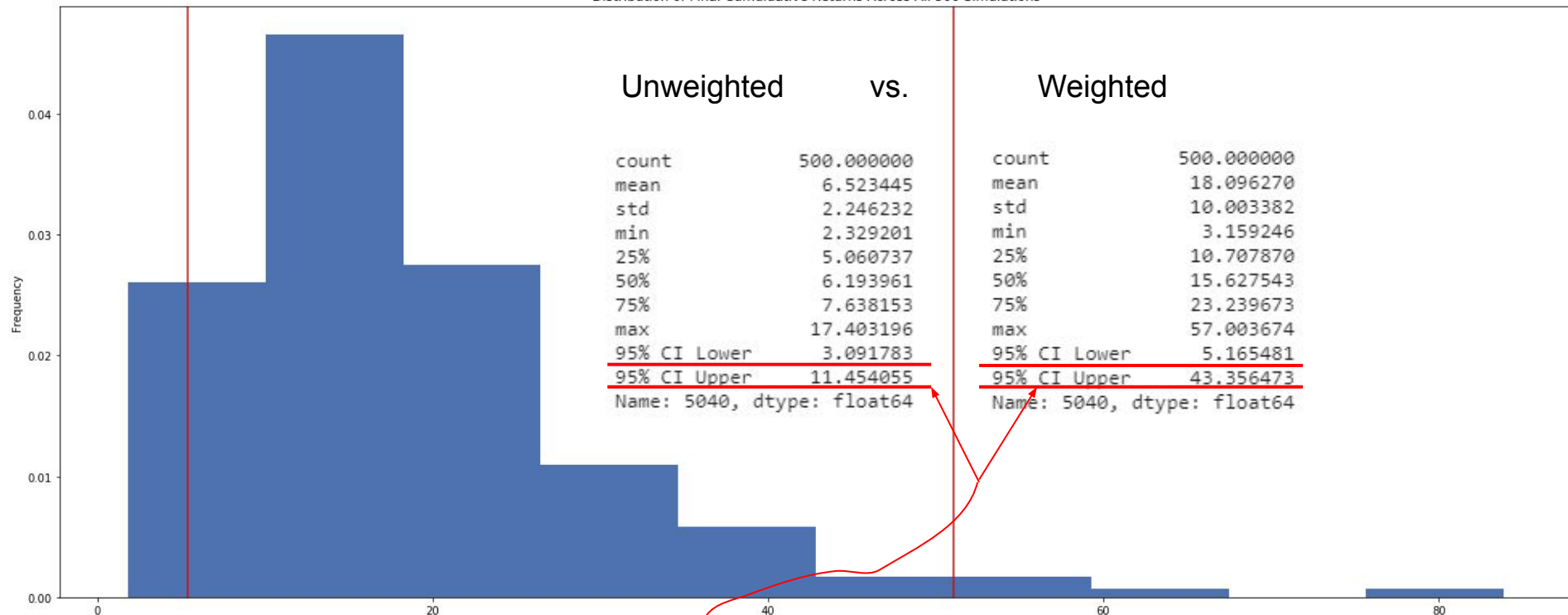


Correlation to S&P 500:

- ❖ Information Technology (.924)
- ❖ Healthcare (.868)
- ❖ Consumer Discretionary (.926)
- ❖ Industrials (.934) (honorable mention)

# Custom Portfolio Simulation

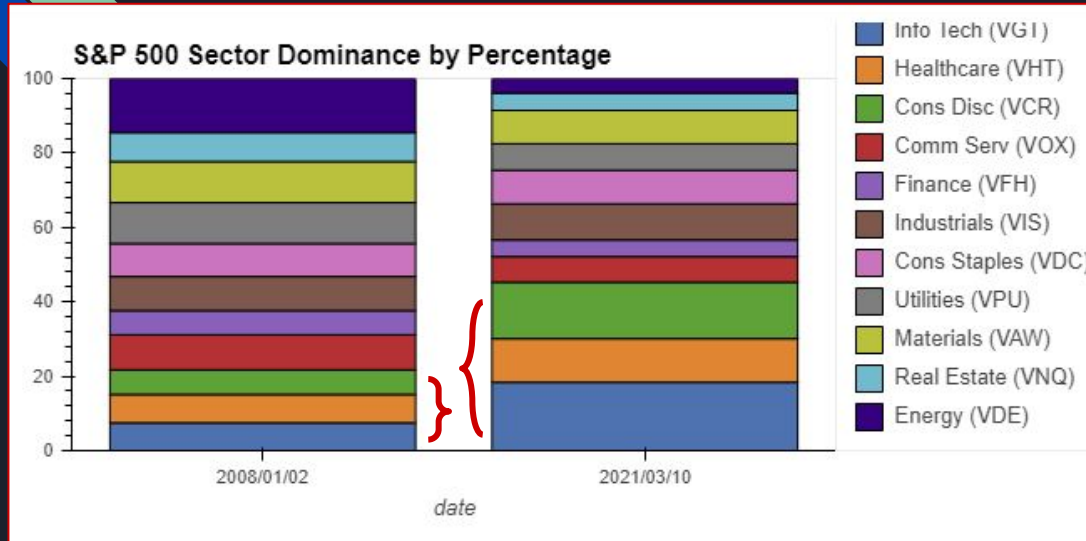
Distribution of Final Cumulative Returns Across All 500 Simulations



Significantly higher projected potential returns



# Conclusions



3 selected sectors grew from ~21% in 2008 to ~45% S&P 500 dominance in 2021.

*"Stocks up after Fed raises economic outlook , but suggest near-zero rates through 2023."* 3/17/2021

Monte Carlo 20 year results with \$20,000, and a confidence level of 95%:

Unweighted: \$61,835.66 - \$229,081.10

Weighted: \$103,309.62 - \$867,129.46

Current corona environment, technology will continue to dominate

Stimulus checks now arriving in bank accounts



# Questions?