1 TASK 1: CAPM MODEL

1.1 INTRODUCTION AND SCOPE

In finance, CAPM model is typically used to identify the excess expected returns of the asset depending on its relation to the market portfolio and risk-free return, or the time value of money. This expected return is then used in pricing the asset.

In this case, we will use CAPM equation to determine the Beta coefficient of assets FORD, GE. MICROSOFT and ORACLE , i.e. how much risk the investment will add to a portfolio that looks like the market – SANDP.

1.2 TOOLS

The regression analysis is done in Jupyter Notebook environment, using python libraries: pandas, numpy, patsy.dmatrices, statsmodels.api

1.3 METHODOLOGY

Simple univariate regression is fit using OLS method. CLRM disturbance terms are checked for each asset. The models are not checked for multicollinearity because they are univariate.

1.4 RESULTS

Asset	Alpha	Beta	R^2	Prob (F-stat)	Interpretation
GE	-0.0087 p = 0.031	1.2977 p = 0.000	0.487	1.61e-29	Industrial companies: high risk involved compared to market index
FORD	-0.0085 p = 0.282	1.8925 p = 0.000	0.339	6.68e-19	
MSFT	0.0018 p = 0.640	1.0128 p = 0.000	0.383	8.22e-22	IT companies: relatively low risk compared to market index
ORACLE	0.0009 p = 0.843	1.0883 p = 0.000	0.331	2.06e-18	

Detailed results of the tests for CLRM disturbance terms, 95% confidence level:

Asset	White Heteroscedasticity	BG Autocorrelation	Jarque-Bera Normality	CUSUM Parameter stability
GE	5.8544 p = 0.054	30.7820 p = 0.002	45.531 p = 1.30e-10	1.2664 p = 0.080
FORD	66.5699 p = 3.50e-15	44.4072 p = 1.30e-05	677.852 p = 6.40e-148	0.7777 p = 0.580
MSFT	0.3081 p = 0.857	16.9580 p = 0.151	19.026 p = 7.39e-05	0.7122 p = 0.691
ORACLE	1.7563 p = 0.415	15.2539 p = 0.227	148.403 p = 5.95e-33	0.8092 p = 0.529

1.5 CONCLUSIONS

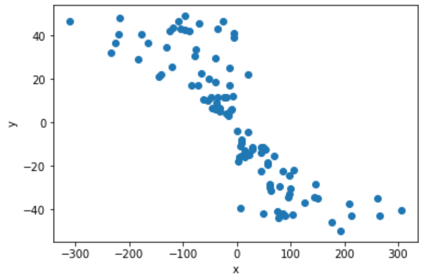
We managed to draw conclusions about risk profile of the given companies. However, in cases of GE and FORD the significance of the parameters might be subject to first type error (false positive), because the residuals bear heteroscedasticity and autocorrelation.

In none of the models the residuals are normally distributed.

2 Task 2: CLRM Disturbance Terms

In this section of the report, each model bears a description of the process and final metrics. The results of formal tests are presented in table form below.

2.1 Model 1



Based on the scatter plot, the functional dependence is unclear. Both linear and cubic model are expected to render high variance of error with heteroscedasticity. Possibly important regressors are missing from the equation.

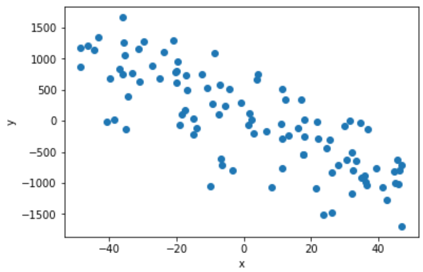
Models tried: linear, cubic

Final equation: $y = 1.1168 - 3.1797x$

$R^2 = 0.726$

Confirmed heteroscedasticity of residuals – significance of coefficients might be inflated

2.2 Model 2

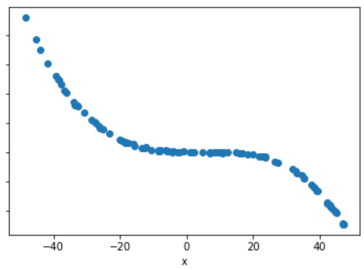


A simple linear model looks like the only option. Log transformation cannot be applied due to negative values of the regressor. High variance of error is expected without any CLRM assumptions violated. Goodness of fit will likely be low.

Final equation: $y = 36.9352 - 22.0048x$

$R^2 = 0.628$

2.3 Model 3



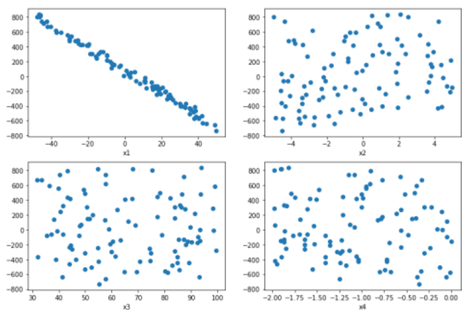
Clear cubic functional dependence. Perfect fit expected.

Models tried: cubic with all regressors, x^1 dropped

Final model: $y = -1571 + 201.6818x2 - 16.0138x3$

$R^2 = 1.000$

2.4 Model 4



At first linear dependence on x1 is suggested. Then, statistically significant coefficients at x2 and x4 made me add these regressors.

Models tried: linear x1, all regressors, without x3

Final model: $y = 14.99 - 15.00x1 + 4.97x2 - 60.09x4$

$R^2 = 1.000$

2.5 GLSR Assumptions

#	White Heteroscedasticity	BG Autocorrelation	Jarque-Bera Normality	CUSUM Parameter stability
1	20.4661 p = 3.596e-05	9.4903 p = 0.660	5.97 p = 0.0505	0.8760 p = 0.426
2	3.2686 p = 0.195	8.3153 p = 0.760	3.999 p = 0.135	0.7647 p = 0.602
3	0.3756 p = 0.995	12.6775 p = 0.392	1.773 p = 0.412	0.8757 p = 0.427
4	8.6235 p = 0.472	6.5735 p = 0.884	0.756 p = 0.685	1.0091 p = 0.260