PROG FOR BUSINESS INTELLIGENCE & ANALYTICS

Final Project

Prepared By:

Abubakr Abdalla, Angad Sharma, Nan Zhang, Wenlan Xu, Yingnan Ji, Zian Zhang

Background

In finance, an exchange rate is the rate at which one currency will be exchanged for another. It is also regarded as the value of one country's currency in relation to another currency. Currencies are traded and priced in pairs. Currency pairs are the quotation of the relative value of a currency unit against the unit of another currency in the foreign exchange market.

In physics, a conservation law states that a particular measurable property of an isolated physical system does not change as the system evolves over time.

A conservation law in volatility and fractal dimension is developed in this project. With the conservation law in exchange rate in hand, people can make money trading exchanges. The base currency can be thought of as a short position because people sell the base currency to purchase the quoted currency. In turn, the quoted currency can be seen as a long position on the currency pair.
[https://www.investopedia.com/ask/answers/06/makingmoneytradingcurrency.asp]

Data Preparation & Cleaning

The preparation & cleaning of the data included sorting the dataset by date (oldest to newest), this was followed up by calculating the fractal dimension & volatility of days 0-100, 1-101, 2-102 & so on, the fractal dimension & volatility of the first 100 days would correspond to the 101st day, the fractal dimension & volatility of days 1-101 would correspond to the 102nd day & so on. The final dataframe which was used to test the model consisted of the date, price, fractal dimension & volatility.

The code below demonstrates the calculation of the fractal dimension for a batch of 100 rows at a time.

```
# calculating fractal dimension for rows 0:100, 1:101, 2,102...
f=[]
for i in range(len(price_df)-100):
    f.append((frac_dim(price_df.iloc[ i:(i+100) , :])))
```

The loop makes use of a frac_dim function defined earlier in the code, the results are appended to a list, the list is then converted to a dataframe. This dataframe is then joined with the larger dataframe (price_df) using the index. The logic behind using the join function is that the dataframe with the fractal dimension values (v_df) will always have a shorter length (100 less) than the larger dataframe containing the prices (price_df).

```python
# joining dataframes on index
price_df = price_df.join(v_df, on = price_df.index)
price_df = price_df.join(f_df, on = price_df.index)
```

	Date	Price
0	2016-03-11	1.1153
1	2016-03-14	1.1106
2	2016-03-15	1.1109
3	2016-03-16	1.1225

+

	frac_dim
0	1.575675
1	1.596856
2	1.599815
3	1.591599

=

	Date	Price	vol	frac_dim
0	2016-03-11	1.1153	0.011851	1.575675
1	2016-03-14	1.1106	0.011841	1.596856
2	2016-03-15	1.1109	0.011801	1.599815
3	2016-03-16	1.1225	0.011745	1.591599

In order to match the fractal dimension & volatility with their corresponding prices the values are shifted down by 100 spots, this is done by using the shift function in the code shown below. The output beneath the code shows the effect on the dataframe.

```python
# shifting volatility & fractal dimension by 100 days
price_df['vol'] = price_df['vol'].shift(periods=100)
price_df['frac_dim'] = price_df['frac_dim'].shift(periods=100)
```

	Date	Price	vol	frac_dim
0	2016-03-11	1.1153	0.011851	1.575675
1	2016-03-14	1.1106	0.011841	1.596856
2	2016-03-15	1.1109	0.011801	1.599815
3	2016-03-16	1.1225	0.011745	1.591599
1040	2020-03-06	1.1286	NaN	NaN
1041	2020-03-09	1.1449	NaN	NaN
1042	2020-03-10	1.1314	NaN	NaN
1043	2020-03-11	1.1279	NaN	NaN

1044 rows × 4 columns

Before

	Date	Price	vol	frac_dim
0	2016-03-11	1.1153	NaN	NaN
1	2016-03-14	1.1106	NaN	NaN
2	2016-03-15	1.1109	NaN	NaN
3	2016-03-16	1.1225	NaN	NaN
1040	2020-03-06	1.1286	0.009032	1.565603
1041	2020-03-09	1.1449	0.009206	1.532216
1042	2020-03-10	1.1314	0.009820	1.467463
1043	2020-03-11	1.1279	0.010061	1.487610

1044 rows × 4 columns

After

The last step to prepare the data is to remove rows that have null values in them, the dropna function is used to drop the first 100 rows of data.

Data Overview

1	Currency Pair	Average Exchange Rate	Average Volatility	Average Fractal Dimension
2	EUR / USD	1.13	0.014	1.47
3	BRL / USD	0.29	0.029	1.45
4	NZD / USD	0.69	0.02	1.45
5	GBP / USD	1.3	0.02	1.45
6	FJD / USD	0.48	0.01	1.52
7	USD / KWD	0.3	0.002	1.5
8	INR / USD	1.14	0.014	1.47
9	USD / CAD	1.31	0.013	1.5
10	USD / JPY	109.78	0.018	1.45
11	USD / CNY	6.77	0.012	1.34

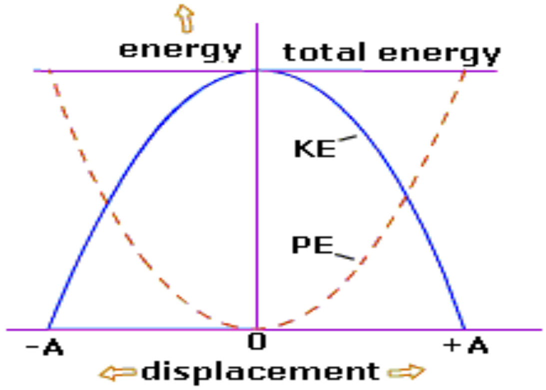
In our database, there are a total ten pairs of currencies from Mar.11.2016 to Mar.11.2020, including 1044 rows in each. The currencies' prices and exchange rate are changing in different ranges. For example, the most volatile currency pair is BRL to USD, the average absolute change is 0.68%; the least volatile currency pair is USD to KWD, the average rate is 0.0431%. The average change in currency pairs will influence the average volatility positively. From the data overview, we can also find out that fractal dimension has a bigger amount than volatility in the quantitative aspect.

Law of Conservation of Energy

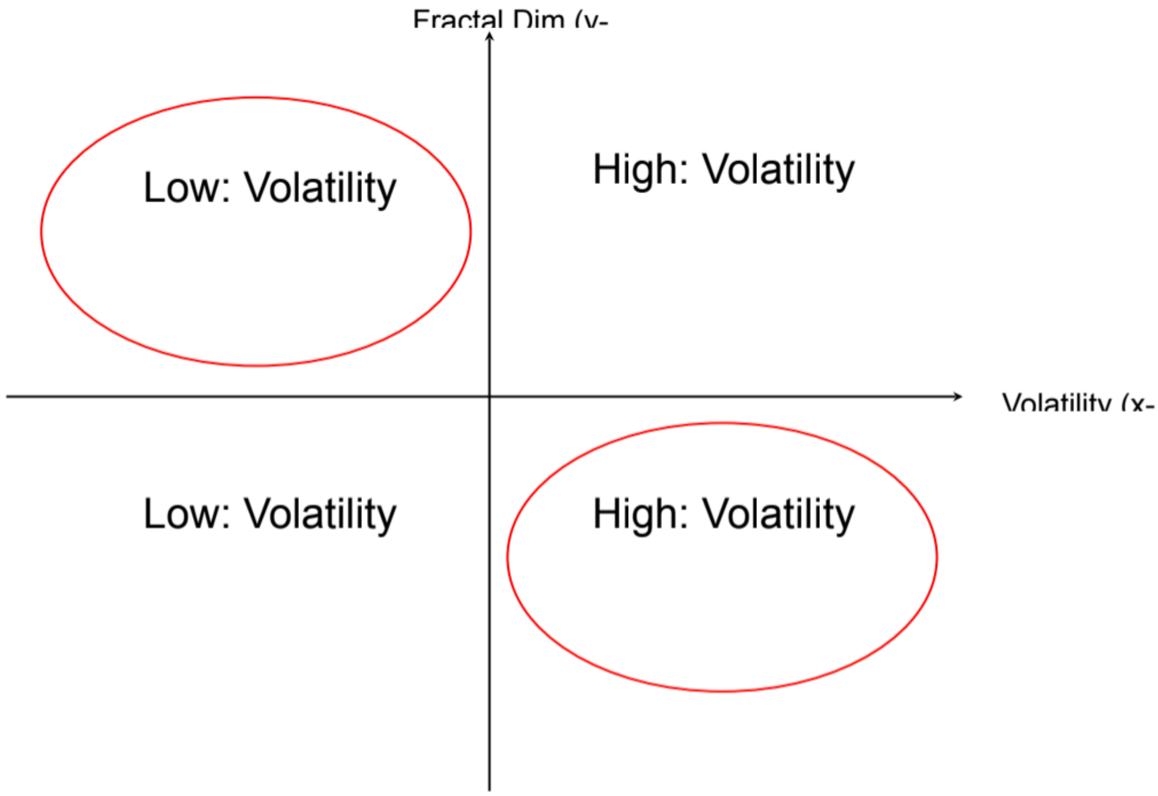
The **law of conservation of energy** states that energy can neither be created nor destroyed - only converted from one form of energy to another.

- The total energy in the system is constant unless new energy is added to the system.
- Mechanical energy takes two forms which are Potential and Kinetic energy.
- Potential energy is determined by the body position or elevation.
- Kinetic energy is determined by the body's motion.
- Kinetic + Potential = Constant (total energy is constant)
- The highest point for the KE is lowest point for PE and vise versa

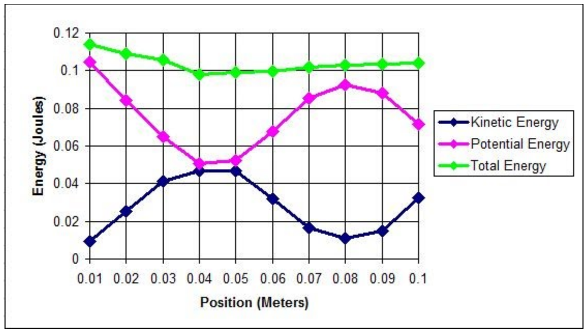
The Graph shows the relationship between KE & PE

Source:  Topper.com

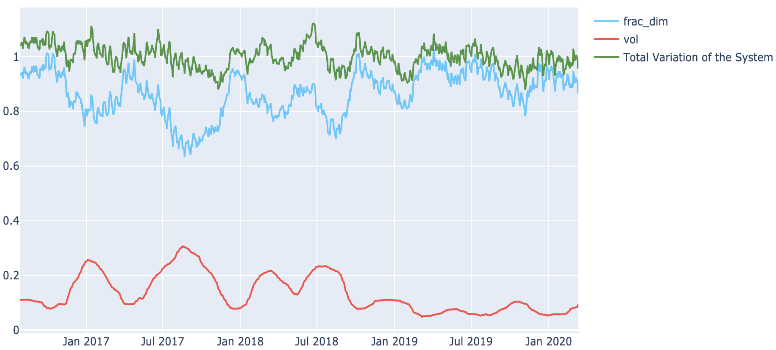
Since we are studying our two parameters, the fractal dimension and the volatility in a similar behavior to the two mechanical energies, using the conservation law would be a good idea to fit the parameters in a model.



Since the other two quadrants are not normal market behavior, the use of a model where the sum of the our parameter is constant.
To validate our conclusion the graph below shows the behavior of the fractal dimension and volatility in a time series, side by side with the behavior of the kinetic and potential energies relative to the position.



We can see the similarities between the energies and our parameters and the form of energies. The constant total



energy is similar to total variation.

Our model has to take the form of:
KE + PE = C
Vol + Frac = C

$$\mathbf{Vol}*\alpha + \mathbf{Fra}*\beta = \mathbf{C}$$

We will try a couple of models and validate them to find the best fit for the data.

Model Design (EUR/USD)

In this project, the problem is to find appropriate coefficients A and B for 10 currency pairs. We will use USR/USD pair as an example to design, validate and calculate A and B for the following equation:

$$A * vol + B * frac.dim = 1$$

By rearranging this equation to the form that is more commonly used:

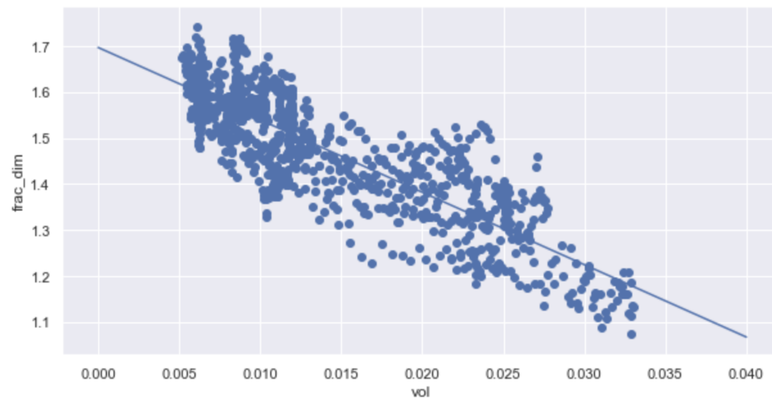
$$frac.dim = -(A/B) * vol + (1/B)$$

This problem becomes nothing but a linear regression problem. In this model, the fractal dimension is the dependent variable and volatility is the independent variable. Three different linear regression models are used in this project: ordinary least square, Ridge regression, and Bayesian regression.

Ordinary Least Square Linear Regression:

In the first model, the coefficient is -15.712 and the intercept is 1.696. These indicate that the fractal dimension is negatively related to volatility: as volatility increased one unit, the fractal dimension decreased 15.712 units. And the r-square of this model is 0.716.

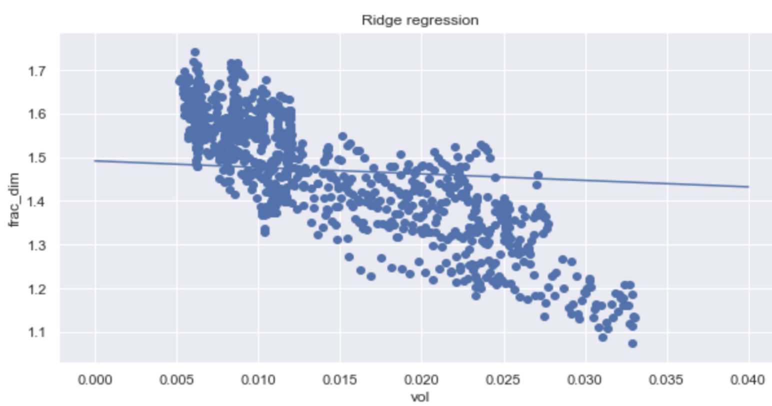
$$frac.dim = -15.712 * vol + 1.6955$$



Ridge Regression:

In the second model, the coefficient is -1.486 and the intercept is 1.491. These indicate that the fractal dimension is negatively related to volatility: as volatility increased one unit, the fractal dimension decreased 1.486 units. And the r-square of this model is 0.129.

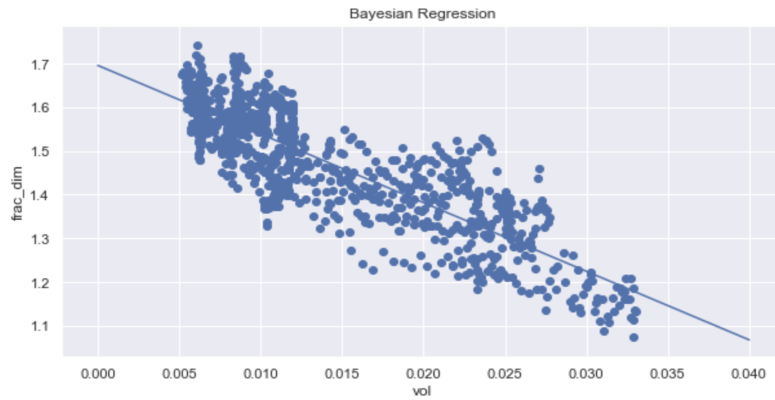
$$frac.dim = -1.4861 * vol + 1.4913$$



Bayesian Regression:

In the third model, the coefficient is -15.705 and the intercept is 1.6954. These indicate that the fractal dimension is negatively related to volatility: as volatility increased one unit, the fractal dimension decreased 1.486 units. And the r-square of this model is 0.716.

$$frac.dim = -15.7054 * vol + 1.6954$$



Model Validation (EUR/USD)

Dep. Variable:	y	**R-squared:**	0.716
Model:	OLS	**Adj. R-squared:**	0.716
Method:	Least Squares	**F-statistic:**	2375.
Date:	Wed, 11 Mar 2020	**Prob (F-statistic):**	9.21e-260
Time:	18:41:28	**Log-Likelihood:**	1123.5
No. Observations:	944	**AIC:**	-2243.
Df Residuals:	942	**BIC:**	-2233.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	1.6955	0.005	325.269	0.000	1.685	1.706
x1	-15.7120	0.322	-48.739	0.000	-16.345	-15.079

Omnibus:	8.031	**Durbin-Watson:**	0.151
Prob(Omnibus):	0.018	**Jarque-Bera (JB):**	5.860
Skew:	-0.062	**Prob(JB):**	0.0534
Kurtosis:	2.634	**Cond. No.**	134.

In the model validation part, we will use three common methods to check your regression.

First, we will use the R-squared to test our regression model. R-squared is a statistical measure to test how close the data are fitted the regression line. From our linear regression output, the R-squared is 0.716, which means that 71.6% of the total points in our database can be described by our predict model. Second, P-value can reflect the relationship between predictor variables and the response variable. In our linear regression output, the P-value are both 0.000, which is less than the common alpha level of 0.05. So we can reject the null hypothesis and prove our predictor variable is significant. The last output we tested is AIC, the method to penalize the inclusion of additional variables to the model. The lower the AIC, the better the model. In our regression model, the AIC is -2243, which is a good and effective for the regression model.

Correlation Calculation (EUR/USD)

As the conclusion of the model validation mentioned, both the ordinary least square linear regression model and the Bayesian regression accurately find the correlation between the fractal dimension and the volatility and the results of these two models should be very similar. For the convenience of calculation, we choose the results of ordinary least square linear regression to calculate A and B in the equation:

$$A * vol + B * frac.dim = 1$$

The correlation between the fractal dimension and the volatility from the ordinary least square linear regression is:

$$frac.dim = -15.712 * vol + 1.6955$$

The equation above is equivalent to:

$$frac.dim = -(A/B) * vol + (1/B)$$

By combining these two equations, we can have two new equations:

$$1/B = 1.6995$$

$$-(A/B) = -15.712$$

Solve these two equations, we have A = 9.2668 and B = 0.5898. So our final equation for "conservative law" between fractal dimension and the volatility is

$$9.2668 * vol + 0.5898 * frac.dim \approx 1$$