**Code explanation**

**Correlation part:**

***corrResult.py***

Step1: get all sector indices

Step2: calculate daily log return

Step3: calculate daily S&P500 log return

Step4: reduce market effect: daily log return of sector indices minus daily S&P500 log return

Step5: calculate correlation matrix of sector indices

**Inference Part:**

***main.py***

Step1: load data: rating indices, sector indices, bundled indices (index in specific sector with specific rating)

Step2: Calculate return of 3 types of indices

Step3: preprocessing: ﻿standardization + ﻿removing outliers

Step4: running regression and estimating coefficients

**regAnalyst package**

***corr.py***

class Corr: help get correlation matrix and visualize it with heatmap

***evaluation.py***

class Kalman: help evaluate the robustness of the regression model with Kalman filter

class rollingReg: help running regression in different time windows, the length of the time window is from start to end with the fixed step

***preprocess.py***

class ﻿Distribution: help analysis distribution based on QQplot, skewness, kurtosis and normal test

class Scale: minmax transformation

class Regular: normalize the data

class ﻿Outlier: draw back outliers to mean + a\*standard deviation

class Missing: fill the missing data with ffill or bfill, or using interpolate method

Any questions please contact Eric Yuan,

Email: eric@bam.money