

# Individual Project

## BA870

Risheng Guo



01

# RISK EXPOSURES

You can enter a subtitle here if you need it

# Data Cleaning(drop)



Sample size

114715 rows



114629 rows



Abnormal Values

Ticker: 62 null  
RET: 15 null & str data type



0

# FF Factor Regression

	TICKER	mktrf	smb	hml
0	A	1.014152	-0.253674	-0.143608
1	AA	1.984149	0.527862	1.924844
2	AAL	1.315550	0.612825	1.248123
3	AAN	1.648965	0.242525	0.815613
4	AAON	0.516779	0.422130	-0.117068
...	...	...	...	...
1918	ZG	1.280585	2.183961	-0.578963
1919	ZION	1.084739	0.869374	1.151468
1920	ZNGA	0.101034	1.200987	-0.714916
1921	ZTS	0.728418	-0.560814	-0.179065
1922	ZUMZ	1.249317	2.336902	0.499124
1923 rows × 4 columns				

## Inner Merge with Ticker.csv

[illegible]

# Variable



mktrf

Market Factor



smb

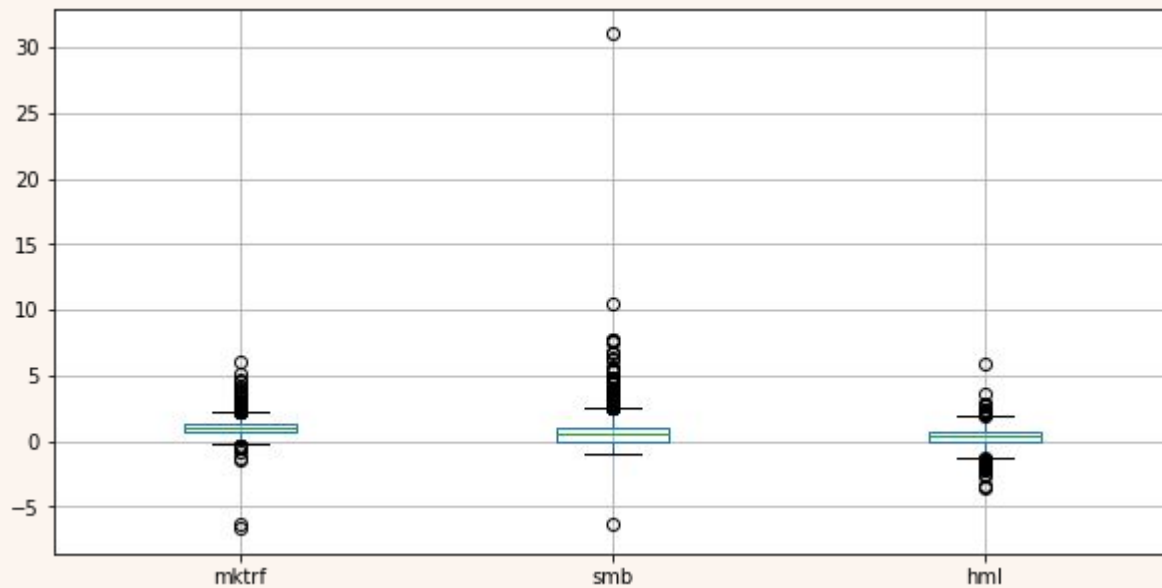
Size Factor



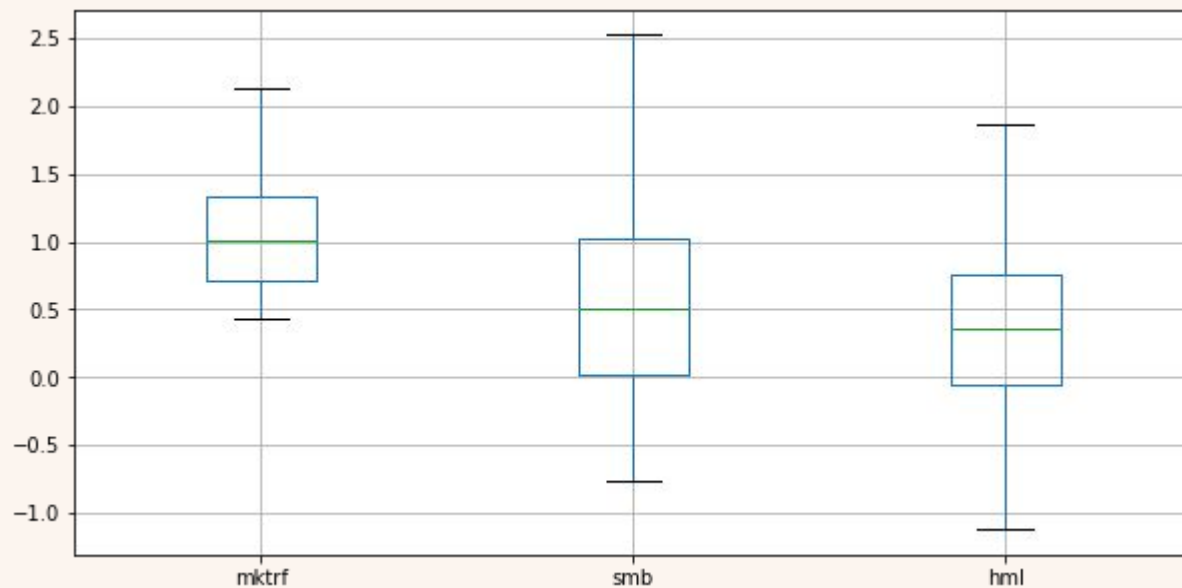
hml

Value Factor

# Outliers



# After Winsorize







02

# Ratios

You can enter a subtitle here if you need it

# Data Cleaning(fill median)



Sample size

1886 rows



Abnormal Values

act: 455 null

lct: 455 null

Dt: 237 null

# Sample Size after Cleaning

1882	294524	20211231	2021	INDL	C
1883	312009	20211231	2021	INDL	C
1884	316056	20211231	2021	INDL	C
1885	317264	20210331	2020	INDL	C

1886 rows × 33 columns

# 10 Ratios



P/E



Market/Book



P/EBIT



P/Sales



DAR



Net Margin



ROA



AT



CR



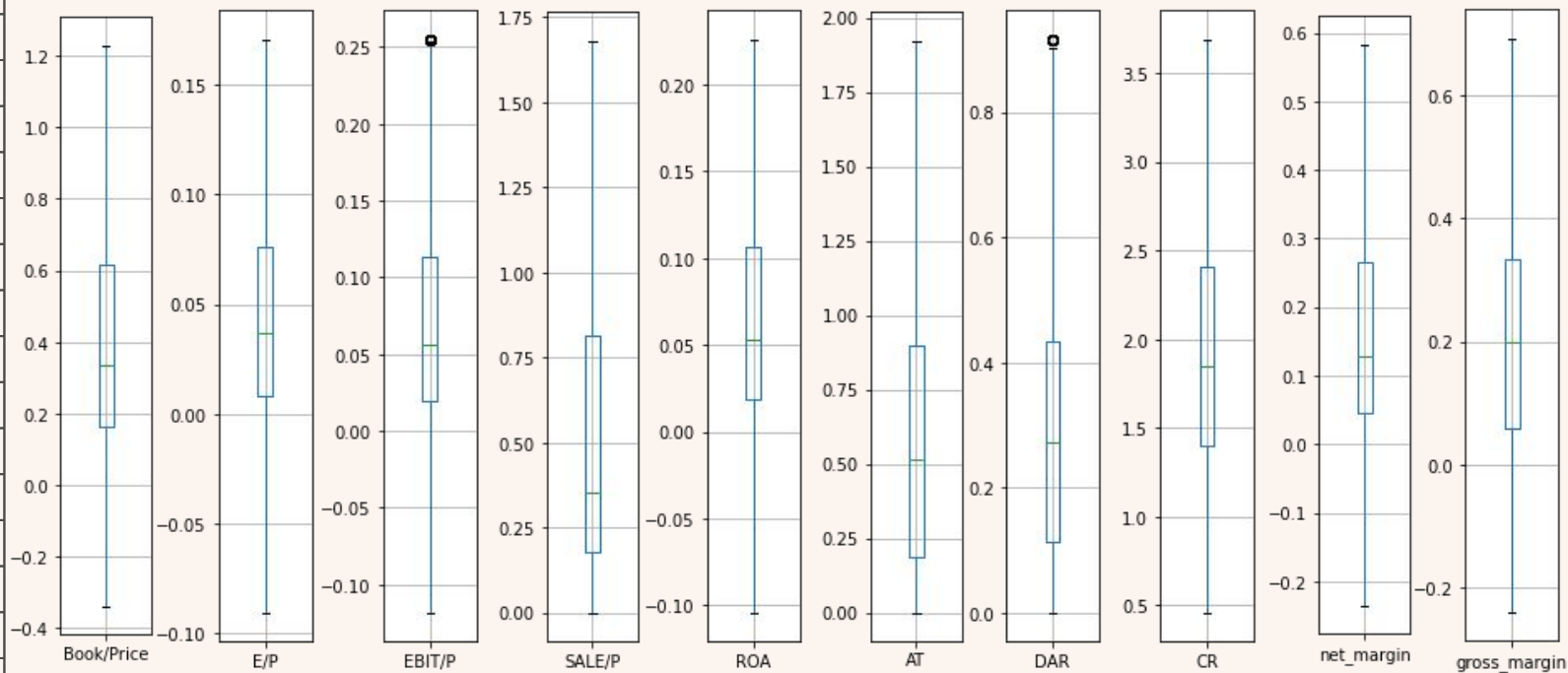
Gross Margin

# 10 Ratios

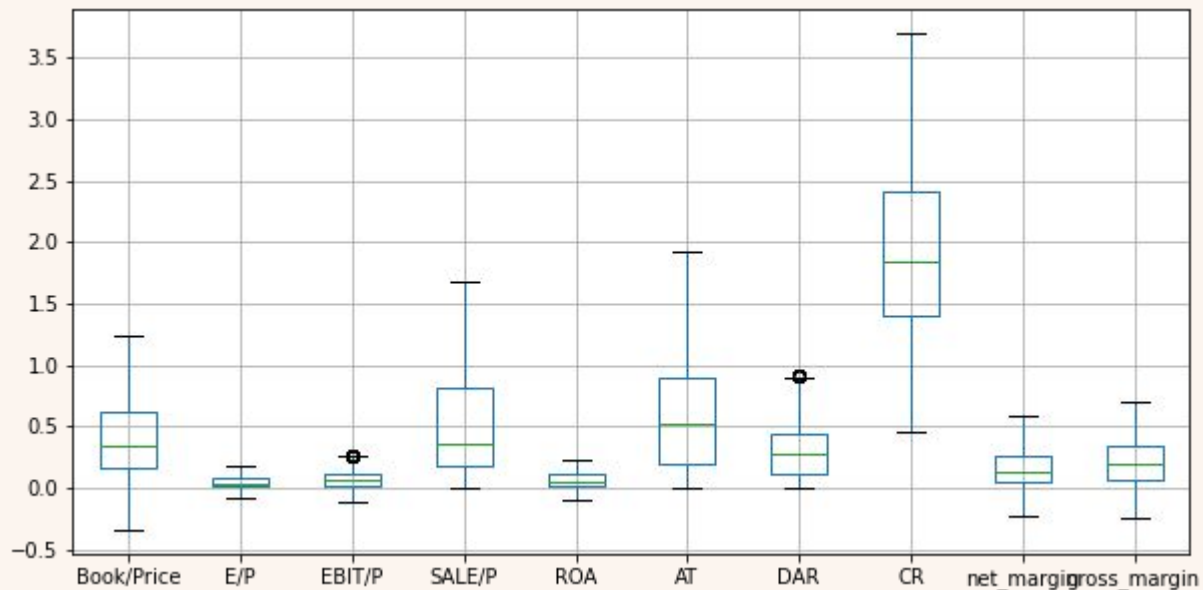
Invert the following ratios to make them "better behaved".

- Book/Price
- E/P
- EBIT/P
- SALE/P

# Outliers



# After Winsorize





03

# Industries

You can enter a subtitle here if you need it



# Industry Data



Sample size

1886 rows



Abnormal Values

None null

# Sample Statistics

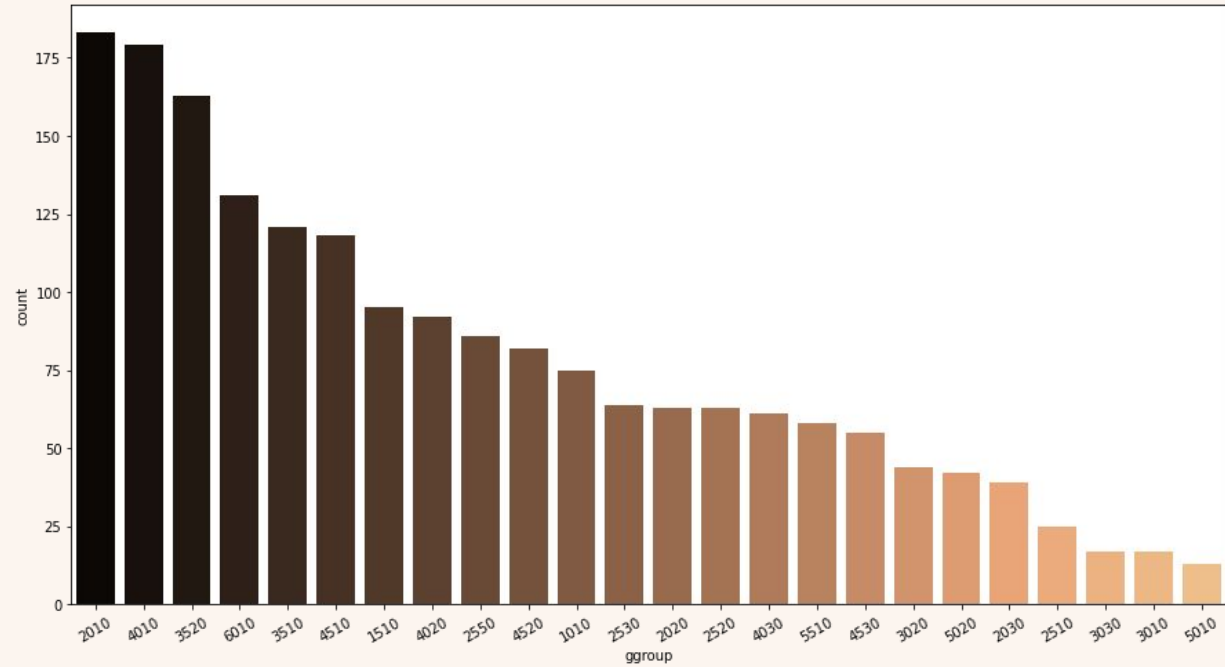


Num of Industries

24

# # stocks in each industry

- 2010(capital goods) has the most stocks(183)
- 5010(Telecommunication Services) has the least stocks(13)



A presentation window with a light gray title bar at the top containing three window control buttons (minimize, maximize, close). The main content area has a light beige background and is framed by a thick border made of a small gray grid pattern. In the center of this area, the number '04' is displayed in a large, bold, black font.

# 04

A light beige rectangular box with a thin gray border, centered horizontally and vertically within the main content area. It contains the word 'Regression' in a bold, black font.

## Regression

You can enter a subtitle here if you need it

# Risk Regression

## OLS Regression Results

Dep. Variable: RetYTD      R-squared: 0.092  
Model: OLS      Adj. R-squared: 0.091  
Method: Least Squares      F-statistic: 63.55  
Date: Thu, 28 Apr 2022      Prob (F-statistic): 3.89e-39  
Time: 00:53:57      Log-Likelihood: 77.279  
No. Observations: 1886      AIC: -146.6  
Df Residuals: 1882      BIC: -124.4  
Df Model: 3

Covariance Type: nonrobust

	coef	std err	t	P> t	[0.025	0.975]
const	-0.0786	0.014	-5.742	0.000	-0.105	-0.052
mktrf	-0.0160	0.012	-1.293	0.196	-0.040	0.008
smb	-0.0113	0.007	-1.591	0.112	-0.025	0.003
hml	0.1192	0.009	13.741	0.000	0.102	0.136

Omnibus: 658.924      Durbin-Watson: 2.003  
Prob(Omnibus): 0.000      Jarque-Bera (JB): 3663.748  
Skew: 1.540      Prob(JB): 0.00  
Kurtosis: 9.094      Cond. No. 5.65




# Financial Characteristics

OLS Regression Results

Dep. Variable:	RetYTD	R-squared:	0.113
Model:	OLS	Adj. R-squared:	0.108
Method:	Least Squares	F-statistic:	23.86
Date:	Wed, 27 Apr 2022	Prob (F-statistic):	9.22e-43
Time:	18:31:38	Log-Likelihood:	99.249
No. Observations:	1886	AIC:	-176.5
Df Residuals:	1875	BIC:	-115.5
Df Model:	10		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	-0.0794	0.025	-3.160	0.002	-0.129	-0.030
Book/Price	0.1713	0.024	7.273	0.000	0.125	0.218
E/P	-0.3811	0.148	-2.567	0.010	-0.672	-0.090
EBIT/P	-0.0973	0.142	-0.686	0.493	-0.376	0.181
SALE/P	0.0933	0.021	4.472	0.000	0.052	0.134
ROA	0.7228	0.124	5.849	0.000	0.480	0.965
AT	-0.1002	0.023	-4.308	0.000	-0.146	-0.055
DAR	0.0611	0.027	2.222	0.026	0.007	0.115
CR	-0.0348	0.007	-5.200	0.000	-0.048	-0.022
net_margin	-0.0776	0.048	-1.623	0.105	-0.171	0.016
gross_margin	-0.0232	0.045	-0.511	0.609	-0.112	0.066



# Industry Dummies

OLS Regression Results

Dep. Variable:	RetYTD	R-squared:	0.326
Model:	OLS	Adj. R-squared:	0.317
Method:	Least Squares	F-statistic:	20.07

	coef	std err	t	P> t	[0.025	0.975]
const	0.5150	0.023	22.140	0.000	0.469	0.561
gggroup_1510	-0.4572	0.031	-14.692	0.000	-0.518	-0.396
gggroup_2010	-0.6056	0.028	-21.924	0.000	-0.660	-0.551
gggroup_2020	-0.5502	0.034	-15.980	0.000	-0.618	-0.483
gggroup_2030	-0.5825	0.040	-14.646	0.000	-0.660	-0.504
gggroup_2510	-0.7230	0.047	-15.540	0.000	-0.814	-0.632
gggroup_2520	-0.7236	0.034	-21.018	0.000	-0.791	-0.656
gggroup_2530	-0.5776	0.034	-16.849	0.000	-0.645	-0.510
gggroup_2550	-0.6663	0.032	-20.932	0.000	-0.729	-0.604
gggroup_3010	-0.4367	0.054	-8.070	0.000	-0.543	-0.331
gggroup_3020	-0.4763	0.038	-12.449	0.000	-0.551	-0.401
gggroup_3030	-0.6688	0.054	-12.359	0.000	-0.775	-0.563
gggroup_3510	-0.5815	0.030	-19.639	0.000	-0.640	-0.523
gggroup_3520	-0.6790	0.028	-24.154	0.000	-0.734	-0.624
gggroup_4010	-0.6029	0.028	-21.756	0.000	-0.657	-0.549
gggroup_4020	-0.6375	0.031	-20.342	0.000	-0.699	-0.576
gggroup_4030	-0.5113	0.035	-14.719	0.000	-0.579	-0.443
gggroup_4510	-0.6237	0.030	-20.963	0.000	-0.682	-0.565
gggroup_4520	-0.6871	0.032	-21.346	0.000	-0.750	-0.624
gggroup_4530	-0.7678	0.036	-21.470	0.000	-0.838	-0.698
gggroup_5010	-0.5032	0.061	-8.314	0.000	-0.622	-0.384
gggroup_5020	-0.5859	0.039	-15.090	0.000	-0.662	-0.510
gggroup_5510	-0.4743	0.035	-13.464	0.000	-0.543	-0.405
gggroup_6010	-0.5509	0.029	-18.884	0.000	-0.608	-0.494

# Combined Regression

## OLS Regression Results

Dep. Variable: RetYTD  
Model: OLS  
Method: Least Squares

R-squared: 0.389  
Adj. R-squared: 0.377  
F-statistic: 22.69

	coef	std err	t	P> t	[0.025	0.975]
const	0.4254	0.037	11.380	0.000	0.352	0.499
mktrf	-0.0293	0.011	-2.685	0.007	-0.051	-0.008
smb	-0.0292	0.007	-3.999	0.000	-0.044	-0.015
hml	0.0693	0.010	6.697	0.000	0.049	0.090
Book/Price	0.0815	0.021	3.868	0.000	0.040	0.123
E/P	-0.0649	0.128	-0.509	0.611	-0.315	0.185
EBIT/P	-0.0740	0.125	-0.594	0.553	-0.319	0.170
SALE/P	0.0706	0.019	3.747	0.000	0.034	0.108
ROA	0.0943	0.132	0.716	0.474	-0.164	0.353
AT	-0.0672	0.021	-3.142	0.002	-0.109	-0.025
DAR	0.0390	0.026	1.479	0.139	-0.013	0.091
CR	0.0020	0.007	0.302	0.763	-0.011	0.015
net_margin	0.0729	0.054	1.353	0.176	-0.033	0.179
gross_margin	0.1129	0.044	2.586	0.010	0.027	0.199

gggroup_1510	-0.4469	0.032	-13.939	0.000	-0.510	-0.384
gggroup_2010	-0.5661	0.029	-19.696	0.000	-0.623	-0.510
gggroup_2020	-0.5264	0.035	-14.893	0.000	-0.596	-0.457
gggroup_2030	-0.5833	0.040	-14.726	0.000	-0.661	-0.506
gggroup_2510	-0.7015	0.046	-15.369	0.000	-0.791	-0.612
gggroup_2520	-0.7016	0.035	-19.812	0.000	-0.771	-0.632
gggroup_2530	-0.5282	0.035	-15.096	0.000	-0.597	-0.460
gggroup_2550	-0.6412	0.034	-18.859	0.000	-0.708	-0.575
gggroup_3010	-0.4219	0.057	-7.444	0.000	-0.533	-0.311
gggroup_3020	-0.4601	0.040	-11.563	0.000	-0.538	-0.382
gggroup_3030	-0.6369	0.054	-11.752	0.000	-0.743	-0.531
gggroup_3510	-0.5027	0.032	-15.564	0.000	-0.566	-0.439
gggroup_3520	-0.5368	0.032	-16.815	0.000	-0.599	-0.474
gggroup_4010	-0.6495	0.036	-18.083	0.000	-0.720	-0.579
gggroup_4020	-0.6449	0.034	-19.193	0.000	-0.711	-0.579
gggroup_4030	-0.5541	0.037	-14.790	0.000	-0.628	-0.481
gggroup_4510	-0.5253	0.033	-16.012	0.000	-0.590	-0.461
gggroup_4520	-0.6324	0.033	-18.971	0.000	-0.698	-0.567
gggroup_4530	-0.6847	0.037	-18.434	0.000	-0.758	-0.612
gggroup_5010	-0.5263	0.060	-8.768	0.000	-0.644	-0.409
gggroup_5020	-0.5455	0.039	-13.938	0.000	-0.622	-0.469
gggroup_5510	-0.4740	0.039	-12.199	0.000	-0.550	-0.398
gggroup_6010	-0.5413	0.031	-17.347	0.000	-0.603	-0.480



# Conclusions



R2 small



Weak explanatory power