

Zeppelin

```
%r
# DS-670 Assignment 2: Data Set Loading and Data Summary
```

FINISHED ▶ ⌵ 📖 ⚙️

```
%r
# Installing packages
install.packages('devtools', repos = 'http://cran.us.r-project.org')
install.packages('knitr', repos = 'http://cran.us.r-project.org')
install.packages('ggplot2', repos = 'http://cran.us.r-project.org')
install.packages(c('devtools','mplot', 'googleVis'), repos = 'http://cran.us.r-project.org')
```

FINISHED ▶ ⌵ 📖 ⚙️

The downloaded binary packages are in
/var/folders/7b/jk0vn7m91g599255yn29jd500000gn/T//RtmpcpHTmj/downloaded_packages

```
%r
# loading the data into RStudio. All the data with the dimension = ARQ.
market_data <- read.csv("/Users/Mohamar/Documents/ARQ_Data_file.csv", header=TRUE, sep=",",
head(market_data)
```

FINISHED ▶ ⌵ 📖 ⚙️

```
ticker dimension calendardate    datekey reportperiod    accoci
1      A      ARQ    2011-03-31 2011-03-09    2011-01-31 -63000000
2      A      ARQ    2011-06-30 2011-06-07    2011-04-30 278000000
3      A      ARQ    2011-09-30 2011-09-07    2011-07-31 271000000
4      A      ARQ    2011-12-31 2011-12-16    2011-10-31 116000000
5      A      ARQ    2012-03-31 2012-03-05    2012-01-31 87000000
6      A      ARQ    2012-06-30 2012-06-04    2012-04-30 68000000

      assets assetsavg  assetsc  assetsnc assetturnover  bvps  capex
1 8.044e+09      NA 4.598e+09 3.446e+09      NA 9.602 -3.8e+07
2 8.649e+09      NA 5.096e+09 3.553e+09      NA 11.392 -5.1e+07
3 8.753e+09      NA 5.223e+09 3.530e+09      NA 12.046 -3.2e+07
4 9.057e+09      NA 5.569e+09 3.488e+09      NA 12.415 -4.9e+07
5 9.099e+09      NA 5.715e+09 3.384e+09      NA 12.920 -4.6e+07
6 9.413e+09      NA 6.010e+09 3.403e+09      NA 13.569 -3.7e+07

      cashneq cashnequsd      cor currentratio  de      debt  debtusd
1 2.655e+09 2.655e+09 7.03e+08      3.270 1.412 2.139e+09 2.139e+09
2 2.975e+09 2.975e+09 7.77e+08      3.201 1.186 2.144e+09 2.144e+09
3 3.101e+09 3.101e+09 7.00e+08      3.170 1.086 2.168e+09 2.168e+09
```



```
%r
# for loop to estimate returns (as a single column) based on price: P(i)/P(i-1)
returns <- vector();
```

FINISHED ▶ ⌵ 📖 ⚙️

```

for(i in 2:length(market_data[,1]))
{
  if(identical(market_data[i,1], market_data[i-1,1]))
  {
    returns[i]=market_data[i,72]/market_data[i-1,72];
  }
  else {
    returns[i]=0;
  }
}

```

```
returns[1]=0;
```

```

# adding the returns column to the market_data
market_data <- cbind(market_data,returns)

```

```

head(market_data)
t(names(market_data))

```

ticker	dimension	calendar	date	key	report	period	accoci	
1	A	ARQ	2011-03-31	2011-03-09	2011-01-31	-63000000		
2	A	ARQ	2011-06-30	2011-06-07	2011-04-30	278000000		
3	A	ARQ	2011-09-30	2011-09-07	2011-07-31	271000000		
4	A	ARQ	2011-12-31	2011-12-16	2011-10-31	116000000		
5	A	ARQ	2012-03-31	2012-03-05	2012-01-31	87000000		
6	A	ARQ	2012-06-30	2012-06-04	2012-04-30	68000000		

	assets	assetsavg	assetssc	assetsnc	assetturnover	bvps	capex
1	8.044e+09	NA	4.598e+09	3.446e+09	NA	9.602	-3.8e+07
2	8.649e+09	NA	5.096e+09	3.553e+09	NA	11.392	-5.1e+07
3	8.753e+09	NA	5.223e+09	3.530e+09	NA	12.046	-3.2e+07
4	9.057e+09	NA	5.569e+09	3.488e+09	NA	12.415	-4.9e+07
5	9.099e+09	NA	5.715e+09	3.384e+09	NA	12.920	-4.6e+07
6	9.413e+09	NA	6.010e+09	3.403e+09	NA	13.569	-3.7e+07

	cashneq	cashneqsd	cor	currentratio	de	debt	debtusd
1	2.655e+09	2.655e+09	7.03e+08	3.270	1.412	2.139e+09	2.139e+09
2	2.975e+09	2.975e+09	7.77e+08	3.201	1.186	2.144e+09	2.144e+09
3	3.101e+09	3.101e+09	7.00e+08	3.170	1.086	2.162e+09	2.162e+09



```
%r
```

FINISHED

```

# selecting my 20 factors
market_data.factors <- market_data[c(1,3,12,17,18,22,26,30,38,40,43,65,67,68,69,70,72,73,74,75,76)]
head(market_data.factors)
t(names(market_data.factors))

```

↓

	ticker	calendardate	bvps	currentratio	de	divyield	ebitdamargin	eps
1	A	2011-03-31	9.602	3.270	1.412	0.000	0.187	0.56
2	A	2011-06-30	11.392	3.201	1.186	0.000	0.205	0.58
3	A	2011-09-30	12.046	3.470	1.086	0.000	0.215	0.95
4	A	2011-12-31	12.415	3.032	1.101	0.000	0.219	0.84
5	A	2012-03-31	12.920	3.352	1.023	0.000	0.210	0.66
6	A	2012-06-30	13.569	3.275	0.993	0.003	0.219	0.73

	evebitda	fcfps	grossmargin	netmargin	payoutratio	pb	pe	pe1
1	9.612	0.236	0.537	0.127	0.000	3.412	14.246	14.322
2	8.655	0.942	0.537	0.119	0.000	3.004	13.343	13.276
3	5.968	0.632	0.527	0.195	0.000	2.115	8.733	8.713
4	5.099	1.329	0.533	0.167	0.000	1.934	8.233	8.197
5	6.399	0.299	0.535	0.141	0.000	2.366	10.141	10.099
6	5.243	0.908	0.530	0.147	0.096	1.992	8.520	8.512

	price	ps	ps1	revenue	sps	tbvps	returns
1	32.940	1.977	1.988	1.519e+09	16.571	17.663	0.0000000
2	34.120	1.929	1.923	1.677e+09	17.741	19.023	1.0358227
3	25.520	1.272	1.275	1.601e+09	18.572	19.284	0.7182122

^

```
%r
# for loop to estimate log of returns (as a single column) based on price: Pt/P(t-1)
logreturns <- vector();
for(i in 2:length(market_data.factors[,1]))
{
  logreturns[i]=log(market_data.factors[i,23]);
}

head(logreturns)

# adding the logreturns column to the market_data
market_data.factors <- cbind(market_data.factors,logreturns)

head(market_data.factors)
t(names(market_data.factors))
```

FINISHED ▶ ⌵ 📖 ⚙

[1] NA 0.03519600 -0.29006866 -0.06451502 0.24573265 -0.12271741

	ticker	calendardate	bvps	currentratio	de	divyield	ebitdamargin	eps
1	A	2011-03-31	9.602	3.270	1.412	0.000	0.187	0.56
2	A	2011-06-30	11.392	3.201	1.186	0.000	0.205	0.58
3	A	2011-09-30	12.046	3.470	1.086	0.000	0.215	0.95
4	A	2011-12-31	12.415	3.032	1.101	0.000	0.219	0.84
5	A	2012-03-31	12.920	3.352	1.023	0.000	0.210	0.66
6	A	2012-06-30	13.569	3.275	0.993	0.003	0.219	0.73

	evebitda	fcfps	grossmargin	netmargin	payoutratio	pb	pe	pe1
1	9.612	0.236	0.537	0.127	0.000	3.412	14.246	14.322
2	8.655	0.942	0.537	0.119	0.000	3.004	13.343	13.276
3	5.968	0.632	0.527	0.195	0.000	2.115	8.733	8.713
4	5.099	1.329	0.533	0.167	0.000	1.934	8.233	8.197
5	6.399	0.299	0.535	0.141	0.000	2.366	10.141	10.099
6	5.243	0.908	0.530	0.147	0.096	1.992	8.520	8.512

	price	ps	ps1	revenue	sps	tbvps	returns	logreturns
1	32.940	1.977	1.988	1.519e+09	16.571	17.663	0.0000000	NA
2	34.120	1.929	1.923	1.677e+09	17.741	19.023	1.0358227	0.03519600

^

```
%r
# replacing all 'Inf' in the data with NA
market_data.factors[mapply(is.infinite, market_data.factors)] <- NA
head(market_data.factors)

# removing all NAs
market_data.factors.noNA <- na.omit(market_data.factors)
head(market_data.factors.noNA)
```

```
ticker calendardate  bvps currentratio  de divyield ebitdamargin  eps
1      A   2011-03-31  9.602           3.270 1.412    0.000      0.187 0.56
2      A   2011-06-30 11.392           3.201 1.186    0.000      0.205 0.58
3      A   2011-09-30 12.046           3.470 1.086    0.000      0.215 0.95
4      A   2011-12-31 12.415           3.032 1.101    0.000      0.219 0.84
5      A   2012-03-31 12.920           3.352 1.023    0.000      0.210 0.66
6      A   2012-06-30 13.569           3.275 0.993    0.003      0.219 0.73
  evebitda fcfps grossmargin netmargin payoutratio  pb  pe  pe1
1    9.612 0.236    0.537    0.127    0.000 3.412 14.246 14.322
2    8.655 0.942    0.537    0.119    0.000 3.004 13.343 13.276
3    5.968 0.632    0.527    0.195    0.000 2.115  8.733  8.713
4    5.099 1.329    0.533    0.167    0.000 1.934  8.233  8.197
5    6.399 0.299    0.535    0.141    0.000 2.366 10.141 10.099
6    5.243 0.908    0.530    0.147    0.096 1.992  8.520  8.512
  price  ps  ps1  revenue  sps  tbvps  returns  logreturns
1 32.940 1.977 1.988 1.519e+09 16.571 17.663 0.00000000 NA
2 34.120 1.929 1.923 1.677e+09 17.741 19.023 1.0358227 0.03519600
3 25.520 1.272 1.275 1.601e+09 12.572 10.281 0.7182122 0.00006866
```



```
%r
# Normalizing all 20 factors

bvps_N <- (market_data.factors.noNA[,3] - mean(market_data.factors.noNA[,3]))/sd(market_data.factors.noNA[,3])
currentratio_N <- (market_data.factors.noNA[,4] - mean(market_data.factors.noNA[,4]))/sd(market_data.factors.noNA[,4])
de_N <- (market_data.factors.noNA[,5] - mean(market_data.factors.noNA[,5]))/sd(market_data.factors.noNA[,5])
divyield_N <- (market_data.factors.noNA[,6] - mean(market_data.factors.noNA[,6]))/sd(market_data.factors.noNA[,6])
ebitdamargin_N <- (market_data.factors.noNA[,7] - mean(market_data.factors.noNA[,7]))/sd(market_data.factors.noNA[,7])
eps_N <- (market_data.factors.noNA[,8] - mean(market_data.factors.noNA[,8]))/sd(market_data.factors.noNA[,8])
evebitda_N <- (market_data.factors.noNA[,9] - mean(market_data.factors.noNA[,9]))/sd(market_data.factors.noNA[,9])
fcfps_N <- (market_data.factors.noNA[,10] - mean(market_data.factors.noNA[,10]))/sd(market_data.factors.noNA[,10])
grossmargin_N <- (market_data.factors.noNA[,11] - mean(market_data.factors.noNA[,11]))/sd(market_data.factors.noNA[,11])
netmargin_N <- (market_data.factors.noNA[,12] - mean(market_data.factors.noNA[,12]))/sd(market_data.factors.noNA[,12])
payoutratio_N <- (market_data.factors.noNA[,13] - mean(market_data.factors.noNA[,13]))/sd(market_data.factors.noNA[,13])
```

```

pb_N <- (market_data.factors.noNA[,14] - mean(market_data.factors.noNA[,14]))/sd(market_da
pe_N <- (market_data.factors.noNA[,15] - mean(market_data.factors.noNA[,15]))/sd(market_da
pe1_N <- (market_data.factors.noNA[,16] - mean(market_data.factors.noNA[,16]))/sd(market_da
price_N <- (market_data.factors.noNA[,17] - mean(market_data.factors.noNA[,17]))/sd(market.
ps_N <- (market_data.factors.noNA[,18] - mean(market_data.factors.noNA[,18]))/sd(market_da
ps1_N <- (market_data.factors.noNA[,19] - mean(market_data.factors.noNA[,19]))/sd(market_da
revenue_N <- (market_data.factors.noNA[,20] - mean(market_data.factors.noNA[,20]))/sd(marki
sps_N <- (market_data.factors.noNA[,21] - mean(market_data.factors.noNA[,21]))/sd(market_da
tbvps_N <- (market_data.factors.noNA[,22] - mean(market_data.factors.noNA[,22]))/sd(market.

# creating a data frame of all the normalized 20 factors
market_data.factors.Norm <- data.frame(bvps_N, currentratio_N, de_N, divyield_N, ebitdamargi
N, pb_N, pe_N, pe1_N, price_N, ps_N, ps1_N, revenue_N, sps_N, tbvps_N)

# cbinding columns 'calendardate', 'returns', and 'logreturns' to the normalized factors.
calendardate <- market_data.factors.noNA$calendardate
returns <- market_data.factors.noNA$returns
logreturns <- market_data.factors.noNA$logreturns
ticker <- market_data.factors.noNA$ticker

market_data.factors.Norm_complete <- cbind(calendardate, market_data.factors.Norm, returns
head(market_data.factors.Norm_complete)

```

	calendardate	bvps_N	currentratio_N	de_N	divyield_N
1	2011-06-30	-0.002459223	0.05235115	0.0017163731	-0.09739857
2	2011-09-30	0.006320483	0.11527403	0.0009796611	-0.09739857
3	2011-12-31	0.011274170	0.01281968	0.0010901679	-0.09739857
4	2012-03-31	0.018053607	0.08767217	0.0005155326	-0.09739857
5	2012-06-30	0.026766189	0.06966079	0.0002945190	-0.07909305
6	2012-09-30	0.032471656	-0.22507090	0.0003681902	-0.06688936

	ebitdamargin_N	eps_N	evebitda_N	fcfps_N	grossmargin_N
1	0.008747055	-0.003881792	0.004304101	0.018521000	0.03571064
2	0.008748785	-0.003630736	0.003810672	0.011472763	0.03331195
3	0.008749477	-0.003705374	0.003651093	0.027319929	0.03475116
4	0.008747920	-0.003827510	0.003889819	0.003901592	0.03523090
5	0.008749477	-0.003780013	0.003677537	0.017747968	0.03403156
6	0.008746363	-0.003800369	0.003909835	0.009585654	0.03091326

	netmargin_N	payoutratio_N	pb_N	pe_N	pe1_N
1	0.008862848	-0.04068744	-0.0009749304	0.0039505328	6.077121e-03
2	0.008875600	-0.04068744	-0.0041137844	0.0001323263	3.477839e-04
3	0.008870002	0.04068744	0.0017528527	0.0002817050	3.001007e-04



```
%r
```

FINISHED

```

# finding out how many different date are there in the data
date <- split(market_data.factors.Norm_complete, as.factor(market_data.factors.Norm_comple
length(date) # there are 20 different calendar dates

```

```
%r
# extracting all data for a particular calendar date and assign each to a dataframe
funct3 <- split(market_data.factors.Norm_complete, market_data.factors.Norm_complete$calendar_date)
factor_names <- c("n.market_data.factors_2011_03_31", "n.market_data.factors_2011_06_30", "n.market_data.factors_2011_09_30", "n.market_data.factors_2011_12_31", "n.market_data.factors_2012_03_31", "n.market_data.factors_2012_06_30", "n.market_data.factors_2012_09_30", "n.market_data.factors_2012_12_31", "n.market_data.factors_2013_03_31", "n.market_data.factors_2013_06_30", "n.market_data.factors_2013_09_30", "n.market_data.factors_2013_12_31", "n.market_data.factors_2014_03_31", "n.market_data.factors_2014_06_30", "n.market_data.factors_2014_09_30", "n.market_data.factors_2014_12_31", "n.market_data.factors_2015_03_31", "n.market_data.factors_2015_06_30", "n.market_data.factors_2015_09_30", "n.market_data.factors_2015_12_31")

for (i in 1:length(funct3)) {
  assign(factor_names[i], funct3[[i]])
}
```

FINISHED ▶ ⌕ 📖 ⚙️

```
%r
# Verifying data for few calendar dates
head(n.market_data.factors_2011_09_30)
```

FINISHED ▶ ⌕ 📖 ⚙️

	calendardate	bvps_N	currentratio_N	de_N	divyield_N
2	2011-09-30	0.006320483	0.1152740	0.0009796611	-0.09739857
21	2011-09-30	0.032793847	-0.3773489	0.0035065832	-0.02417646
43	2011-09-30	-0.178724574	-0.6068186	-0.0297117596	-0.09739857
51	2011-09-30	-0.348935417	-0.5172295	-0.0520709679	-0.09739857
78	2011-09-30	-0.125093189	-0.2643685	-0.0036689914	-0.03027830
97	2011-09-30	-0.013131802	-0.4325526	0.0204583256	-0.07299120

	ebitdamargin_N	eps_N	evebitda_N	fcfps_N	grossmargin_N
2	0.008748785	-0.003630736	0.003810672	0.0114727627	0.03331195
21	0.008731313	-0.004166774	0.003704347	0.0006048354	-0.05088193
43	0.008712802	-0.004329621	0.022230964	-0.0030784370	-0.03289178
51	0.008712975	-0.004601033	0.005312260	-0.0563494697	0.01316299
78	0.008738406	-0.004207486	0.005370656	-0.0047836557	-0.04680416
97	0.008737368	-0.003305042	0.003936830	0.0236366562	0.02563615

	netmargin_N	payoutratio_N	pb_N	pe_N	pe1_N
2	0.008875600	-0.040687439	-0.004113784	0.0001323263	0.0003477839
21	0.008847411	-0.001323075	-0.009074516	0.0011734294	0.0017766653
43	0.008822810	0.040687439	0.011021651	0.0000780520	0.0118078026



```
%r
head(n.market_data.factors_2012_03_31)
```

FINISHED ▶ ⌕ 📖 ⚙️

```

calendardate      bvps_N currentratio_N      de_N  divyield_N
4      2012-03-31  0.01805361      0.08767217  0.0005155326 -0.09739857
23     2012-03-31  0.02259113      -0.40331339  0.0050021085 -0.02417646
45     2012-03-31 -0.17967772      -0.61430385 -0.0273174457 -0.09739857
53     2012-03-31 -0.50318438      -0.50576774 -0.0351781624 -0.09739857
80     2012-03-31 -0.12478442      -0.18998379 -0.0035953202 -0.01807461
99     2012-03-31  0.02485989      -0.41875172  0.0160601552 -0.07909305
      ebitdamargin_N      eps_N  evebitda_N      fcfps_N grossmargin_N
4      0.008747920 -0.003827510  0.003889819  0.0039015916  0.03523090
23     0.008729583 -0.004214271  0.004033422 -0.0136735293 -0.05687864
45     0.008717300 -0.004302480  0.004505182 -0.0053520619 -0.02857415
53     0.008669207 -0.007634058  0.001233348  0.0541259669  0.01292312
80     0.008741001 -0.004221057  0.005141295  0.0007639892 -0.04320613
99     0.008736330 -0.003033631  0.003899185  0.0448723133  0.02707537
      netmargin_N payoutratio_N      pb_N      pe_N      pe1_N
4  0.008866539 -0.04068744 -0.003227561  0.001298494  0.002088056
23 0.008845565 -0.04068744 -0.008954469  0.014841946  0.022868362
45 0.008838182 -0.04068744 -0.012068054 -0.000810858 -0.014641171

```

```

%r
head(n.market_data.factors_2013_12_31)

```

FINISHED ▶ ✖ 📖 ⚙

```

11     2013-12-31  0.05835434      0.03106497  0.0005007983 -0.04858383
30     2013-12-31 -0.02344191      -0.42927785  0.0084351863 -0.03027830
60     2013-12-31 -0.23314801      -0.45383883 -0.1284385309 -0.09739857
69     2013-12-31 -0.07041468      -0.22974918 -0.0014146528 -0.09739857
87     2013-12-31 -0.11540061      0.03527542 -0.0047151224 -0.06078752
106    2013-12-31  0.12429610      -0.35886973  0.0126491787 -0.08519489
      ebitdamargin_N      eps_N  evebitda_N      fcfps_N grossmargin_N
11     0.008749823 -0.003847866  0.004587083  0.020726416  0.03379169
30     0.008663152 -0.005761318  0.198262282  0.007607600 -0.05543943
60     0.008686161 -0.012248057 -0.016637689 -0.086838780  0.02083878
69     0.008716262 -0.002171898  0.022653509 -0.023518325 -0.02569573
87     0.008746882 -0.004180345  0.005341458  0.002105428 -0.02713494
106    0.008728718 -0.003813939  0.004632257  0.027956544  0.02635576
      netmargin_N payoutratio_N      pb_N      pe_N      pe1_N
11  0.008863519 -0.01074540 -0.002486100  0.008479373  0.01371749
30  0.008772574 -0.04634083 -0.007485669 -0.011554202 -0.01728095
60  0.008797240 -0.04068744 -0.034093548 -0.014964912 -0.01471448
69  0.008830180  0.04068744  0.007000030  0.203700271  0.21700566

```

```

%r
head(n.market_data.factors_2014_06_30)

```

FINISHED ▶ ✖ 📖 ⚙

132014-06-300.071416500.035275424.403692e-05-0.04248199

322014-06-30-0.02212630-0.400740346.608141e-03-0.05468567

622014-06-30-0.07922124-0.464364966.640705e-02-0.09739857

892014-06-30-0.11201760-0.02647788-4.368868e-03-0.04858383

1082014-06-300.17861214-0.404249051.796824e-02-0.08519489

1272014-06-300.11463037-0.35232013-8.326503e-040.01243460

ebitdamargin_Neps_Nevebitda_Nfcfps_Ngrossmargin_N

130.008742904-0.0039903570.0045589870.0158153860.031393001

320.008731486-0.004193915-0.0250054070.002128164-0.048962979

620.008733216-0.0034611040.008001234-0.0106041360.025156417

890.008749650-0.0041328480.005269656-0.000350087-0.020658485

1080.008734081-0.0029793480.0048381130.0575136680.015321805

1270.008769026-0.0034000360.0045538450.0268879400.001409426

netmargin_Npayoutratio_Npb_Npe_Npe1_N

130.0088563040.004120933-0.0029874690.0088081860.013519108

320.0088469070.011658790-0.005533154-0.014134183-0.020250459

620.008855632-0.0406874390.014652255-0.075430065-0.016454757

890.0088635100.0101386800.0000877620.0117268200.022381052

%r

head(n.market_data.factors_2015_09_30)

FINISHED

calendardatebvps_Ncurrentratio_Nde_Ndivyield_N

182015-09-300.010508970.2572598-0.001370450-0.03027830

372015-09-30-0.02634163-0.30600520.005790390-0.01807461

402015-09-30-0.06983743-0.13758700.001163839-0.09739857

672015-09-30-0.07527440-0.46997890.081008683-0.04248199

742015-09-30-0.01758878-0.1230844-0.002770203-0.09739857

752015-09-30-0.01758878-0.1230844-0.002770203-0.09739857

ebitdamargin_Neps_Nevebitda_Nfcfps_Ngrossmargin_N

180.008746536-0.0040649950.0055059950.0020372190.02827471

370.008729756-0.0042617680.003974292-0.000259142-0.04944272

400.008730275-0.0042007010.006878670-0.0127186070.14676980

670.008742731-0.0025383040.003908917-0.0137644740.05034262

740.008729410-0.0041599890.005588815-0.037273756-0.01730032

750.008729410-0.0041599890.005628480-0.037273756-0.01730032

netmargin_Npayoutratio_Npb_Npe_Npe1_N

180.0088599950.02694389-0.0014480530.0276051920.042517365

370.0088442220.27338993-0.0082236000.0113343310.018834095

400.0088500050.010687110.0017610170.0261511080.026282722

READY