Week 6 Assignment

David Russo

2/18/2017

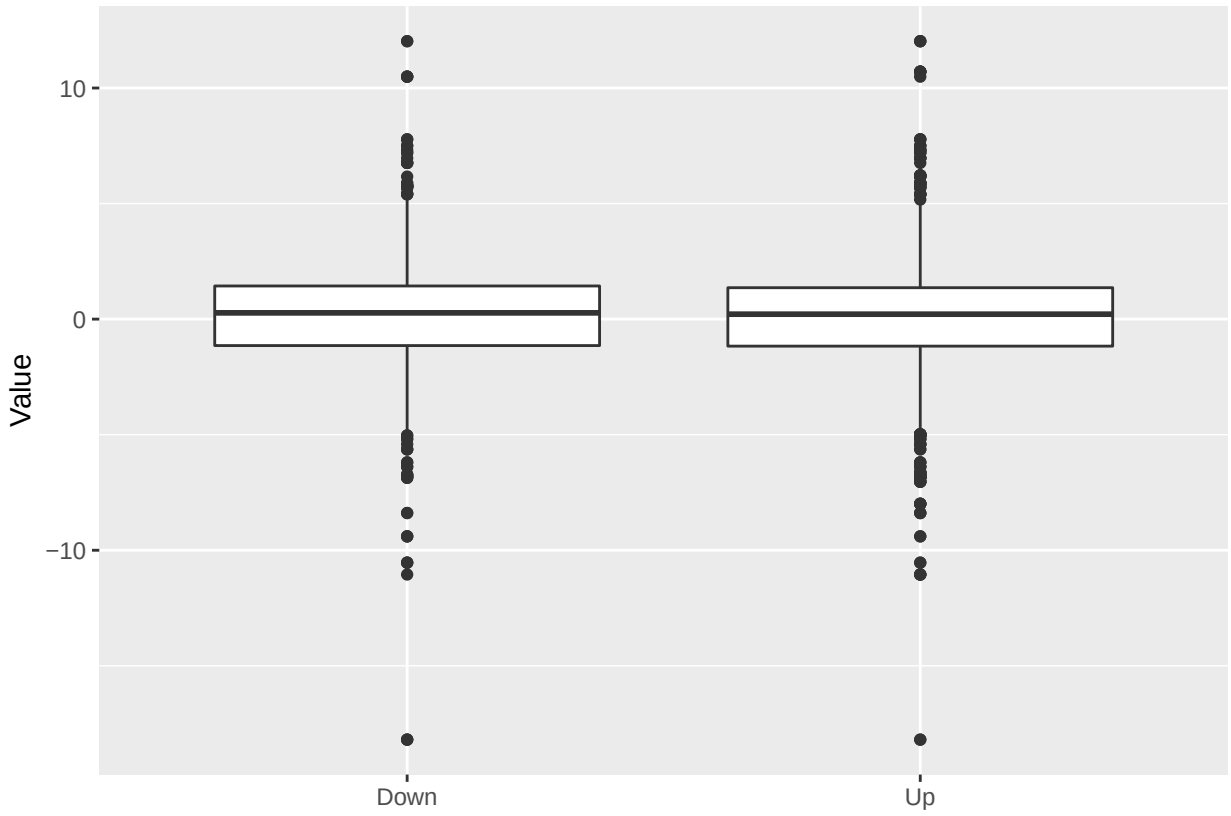
- 4.7 #10
- a.

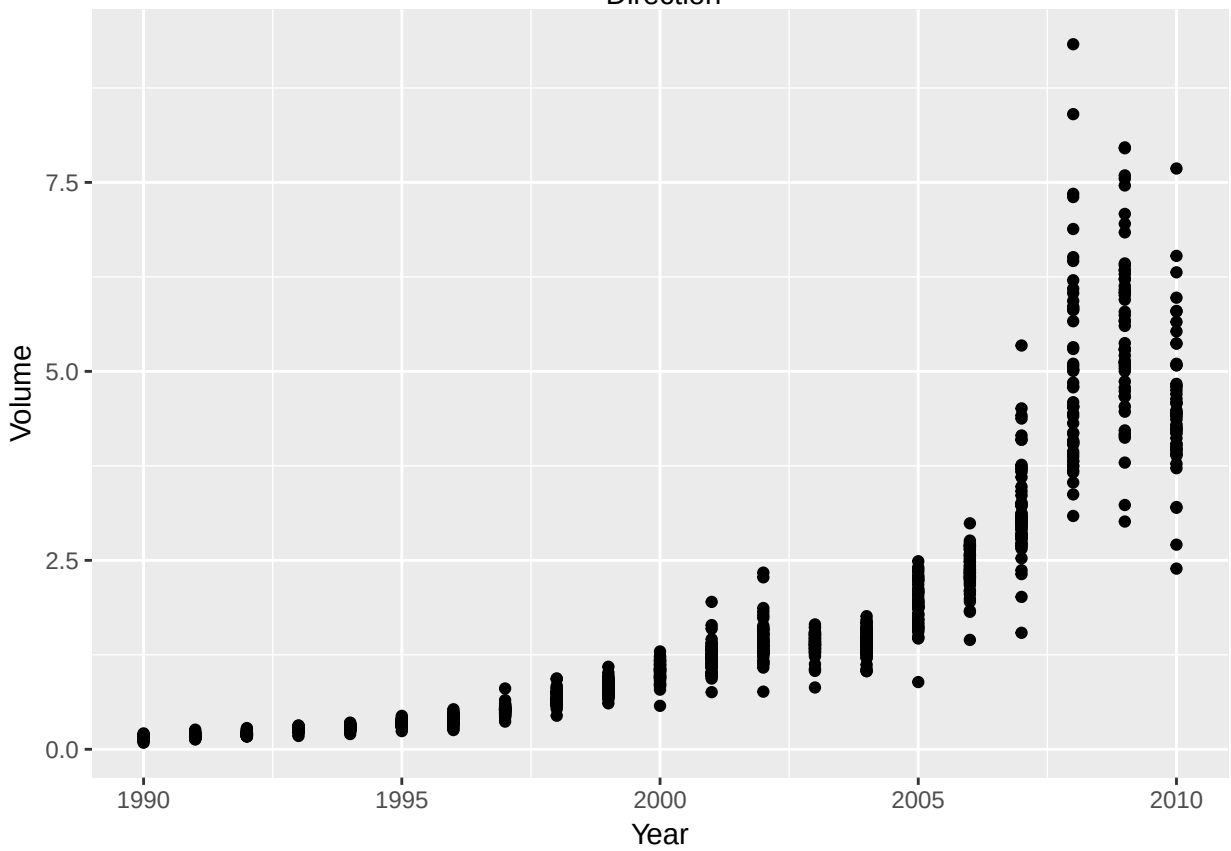
Direction	Count	Percent
Down	484	44%
Up	605	56%

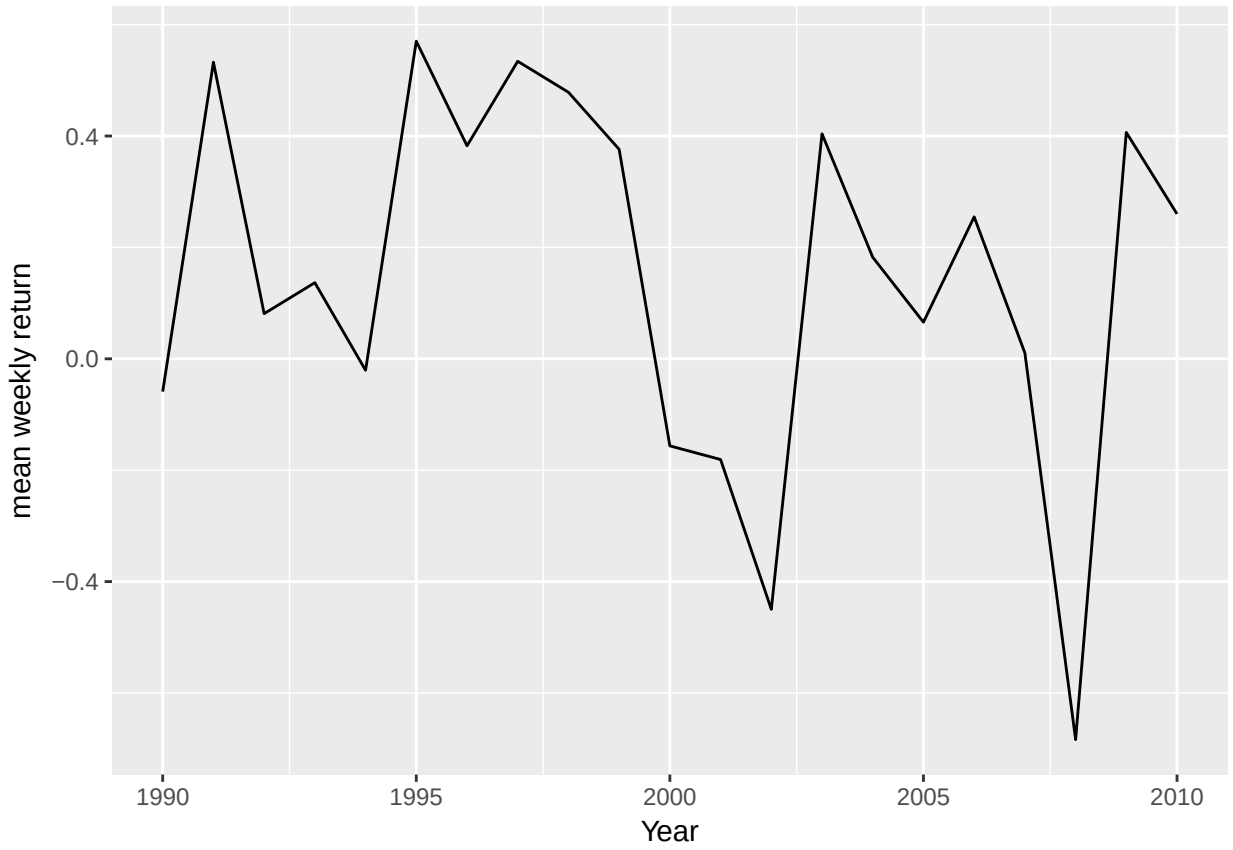
Direction	Mean Lag1	Mean Lag2	Mean Lag3	Mean Lag4	Mean Lag5	Mean Volume	Mean Today
Down	0.2823	-0.0404	0.2076	0.2000	0.1878	1.6085	-1.7466
Up	0.0452	0.3043	0.0989	0.1025	0.1015	1.5475	1.6671

	Year	Lag1	Lag2	Lag3	Lag4	Lag5	Volume	Today
Year	1.0000000	-0.0322893	-0.0333900	-0.0300065	-0.0311279	-0.0305191	0.8419416	-0.0324599
Lag1	-0.0322893	1.0000000	-0.0748531	0.0586357	-0.0712739	-0.0081831	-0.0649513	-0.0750318
Lag2	-0.0333900	-0.0748531	1.0000000	-0.0757209	0.0583815	-0.0724995	-0.0855131	0.0591667
Lag3	-0.0300065	0.0586357	-0.0757209	1.0000000	-0.0753959	0.0606572	-0.0692877	-0.0712436
Lag4	-0.0311279	-0.0712739	0.0583815	-0.0753959	1.0000000	-0.0756750	-0.0610746	-0.0078259
Lag5	-0.0305191	-0.0081831	-0.0724995	0.0606572	-0.0756750	1.0000000	-0.0585174	0.0110127
Volume	0.8419416	-0.0649513	-0.0855131	-0.0692877	-0.0610746	-0.0585174	1.0000000	-0.0330778
Today	-0.0324599	-0.0750318	0.0591667	-0.0712436	-0.0078259	0.0110127	-0.0330778	1.0000000

We see that for the market data, about 56% of the weeks had positive market performance while 44% of the weeks had negative market performance. The value of *today* does not appear to be highly correlated with any of the *lag* or *volume* covariates.


10
0
−10
Value
Down
Up
Direction


7.5
5.0
2.5
0.0
Volume
1990
1995
2000
2005
2010
Year



From the side-by-side boxplots, we see that markets finished up and down with relatively equal magnitudes. Furthermore, we can see that the number of trades has increased exponentially since 1990. Lastly, we can see that the average weekly return has varied from year to year, with several more down years between the years 2000 and 2010 vs 1990 and 2000.

- b.

```
glm.b <- glm(Direction ~ Lag1 + Lag2 + Lag3 + Lag4 + Lag5 + Volume,
             data = Weekly,
             family = binomial)

summary(glm.b)
```

```
##
## Call:
## glm(formula = Direction ~ Lag1 + Lag2 + Lag3 + Lag4 + Lag5 +
##     Volume, family = binomial, data = Weekly)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.6949  -1.2565   0.9913   1.0849   1.4579
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.26686    0.08593   3.106   0.0019 **
## Lag1        -0.04127    0.02641  -1.563   0.1181
## Lag2         0.05844    0.02686   2.175   0.0296 *
## Lag3        -0.01606    0.02666  -0.602   0.5469
```

```
## Lag4        -0.02779    0.02646  -1.050   0.2937
## Lag5        -0.01447    0.02638  -0.549   0.5833
## Volume      -0.02274    0.03690  -0.616   0.5377
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1496.2  on 1088  degrees of freedom
## Residual deviance: 1486.4  on 1082  degrees of freedom
## AIC: 1500.4
##
## Number of Fisher Scoring iterations: 4
```