Combined Data Analysis for Spring 2016 and Fall 2016

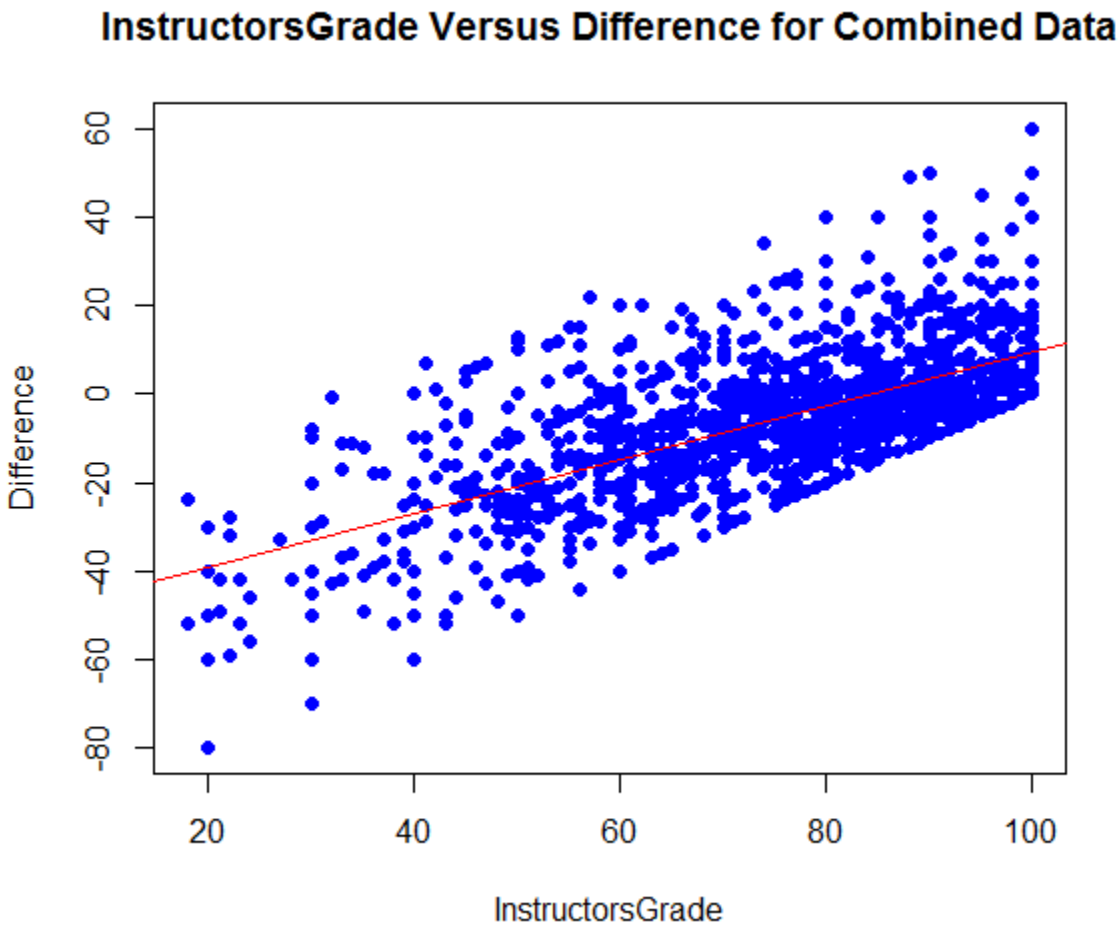
Prepared by Ramjee Sharma

Overall Analysis

Descriptive Statistics

	StudentsGrade	InstructorsGrade	Difference
nbr.val	2567	2567	2567
nbr.null	0	0	331
nbr.na	0	0	0
min	33	18	-80
max	100	100	60
range	67	82	140
sum	209026	197489	-11537
median	82	80	-2
mean	81.43	76.93	-4.49
SE.mean	0.29	0.37	0.34
CI.mean.0.95	0.57	0.72	0.66
var	216.8	350.41	293.56
std.dev	14.72	18.72	17.13
coef.var	0.18	0.24	-3.81

Scatter Plot



T Test

Paired t-test

data: StudentsGrade and InstructorsGrade

t = 10, df = 3000, p-value <0.0000000000000002

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval:

3.9 Inf

sample estimates:

mean of the differences

4.5

Paired t-test

data: StudentsGrade and InstructorsGrade

t = 10, df = 3000, p-value = 1

alternative hypothesis: true difference in means is less than 0

95 percent confidence interval:

-Inf 5.1

sample estimates:

mean of the differences

4.5

Interpretation:

Students significantly over estimated their grades.

Regression Analysis

Residuals:

Min	1Q	Median	3Q	Max
-40.80	-9.57	-1.69	7.16	50.43

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-51.3878	1.0675	-48.1	<0.0000000000000002	***
InstructorsGrade	0.6095	0.0135	45.2	<0.0000000000000002	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13 on 2565 degrees of freedom

Multiple R-squared: 0.443, Adjusted R-squared: 0.443

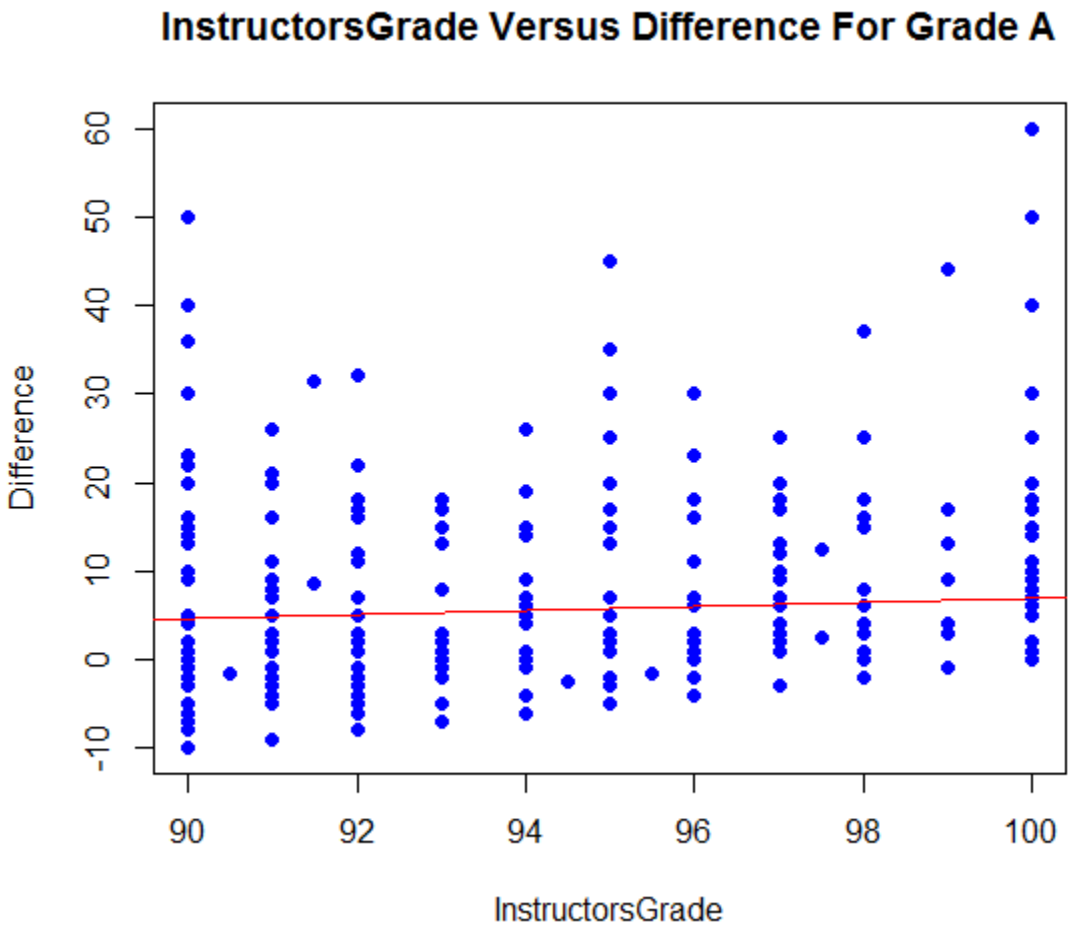
F-statistic: 2.04e+03 on 1 and 2565 DF, p-value: <0.0000000000000002

Grade A

Descriptive Statistics

	StudentsGrade	InstructorsGrade	Difference
nbr.val	849	849	849
nbr.null	0	0	225
nbr.na	0	0	0
min	40	90	-10
max	100	100	60
range	60	10	70
sum	76285	81421.5	5136.5
median	93	96	2
mean	89.85	95.903	6.05
SE.mean	0.39	0.134	0.38
CI.mean.0.95	0.77	0.263	0.75
var	131.04	15.218	122.96
std.dev	11.45	3.901	11.09
coef.var	0.13	0.041	1.83

Scatter Plot



T Test

Paired t-test

data: StudentsGrade and InstructorsGrade

t = -20, df = 800, p-value = 1

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval:

-6.7 Inf

sample estimates:

mean of the differences

-6.1

Paired t-test

data: StudentsGrade and InstructorsGrade

t = -20, df = 800, p-value <0.0000000000000002

alternative hypothesis: true difference in means is less than 0

95 percent confidence interval:

-Inf -5.4

sample estimates:

mean of the differences

-6.1

Interpretation:

Students significantly under estimate their grades.

Regression Analysis

Residuals:

Min	1Q	Median	3Q	Max
-14.67	-7.01	-3.90	5.33	52.99

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-16.4259	9.3426	-1.76	0.079	.
InstructorsGrade	0.2344	0.0973	2.41	0.016	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11 on 847 degrees of freedom

Multiple R-squared: 0.0068, Adjusted R-squared: 0.00563

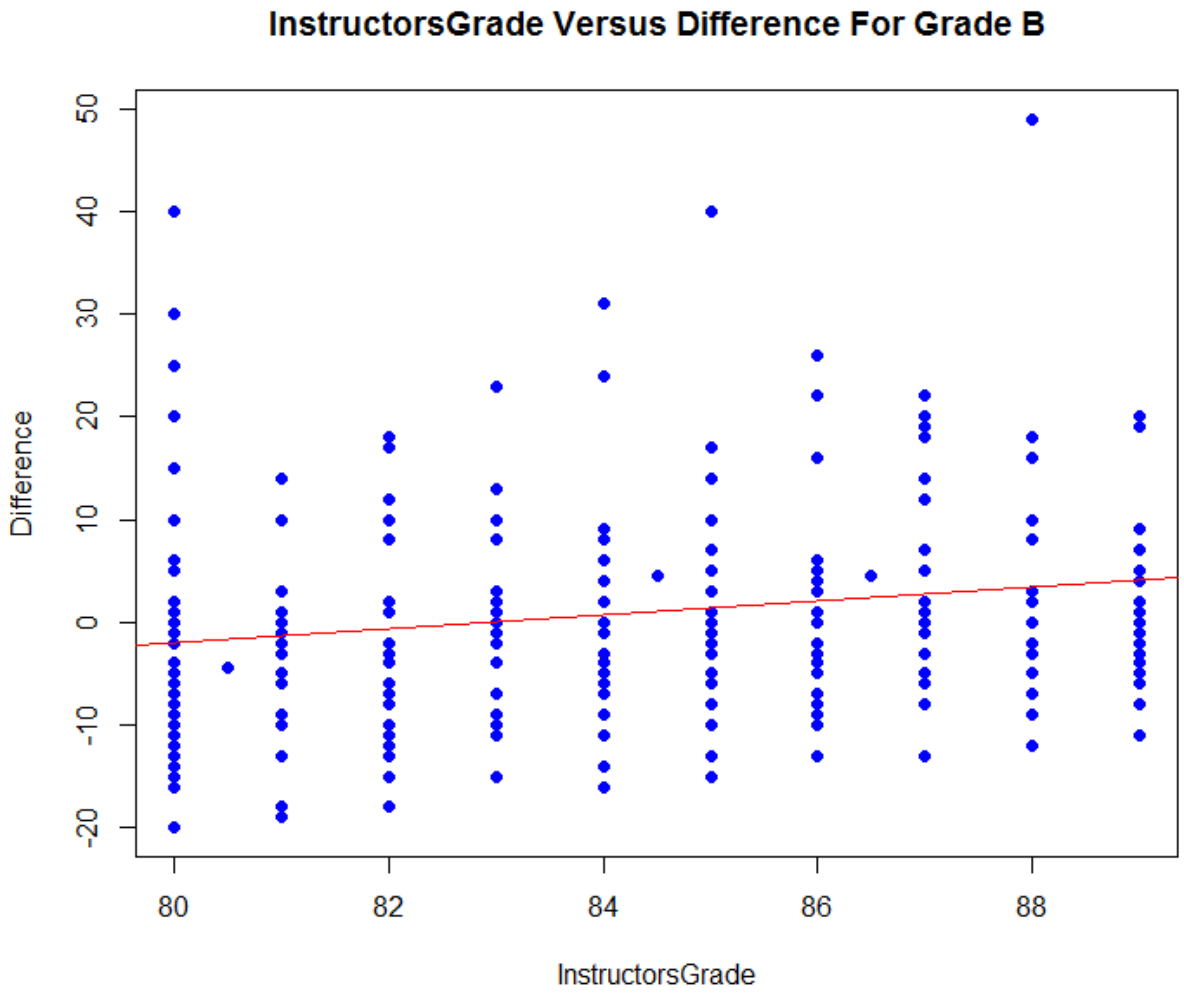
F-statistic: 5.8 on 1 and 847 DF, p-value: 0.0163

Grade B

Descriptive Statistics

	StudentsGrade	InstructorsGrade	Difference
nbr.val	537.00	537.000	537.00
nbr.null	0.00	0.000	51.00
nbr.na	0.00	0.000	0.00
min	39.00	80.000	-20.00
max	100.00	89.000	49.00
range	61.00	9.000	69.00
sum	44616.00	44877.500	261.50
median	85.00	83.000	0.00
mean	83.08	83.571	0.49
SE.mean	0.52	0.134	0.53
CI.mean.0.95	1.02	0.264	1.03
var	145.25	9.664	148.62
std.dev	12.05	3.109	12.19
coef.var	0.15	0.037	25.03

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = -0.9, df = 500, p-value = 0.8
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -1.4  Inf
sample estimates:
mean of the differences
                  -0.49
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = -0.9, df = 500, p-value = 0.2
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
```

```
 -Inf 0.38
sample estimates:
mean of the differences
                  -0.49
```

Interpretation:

Students estimated grades are not significantly different from their actual grades.

Regression Analysis

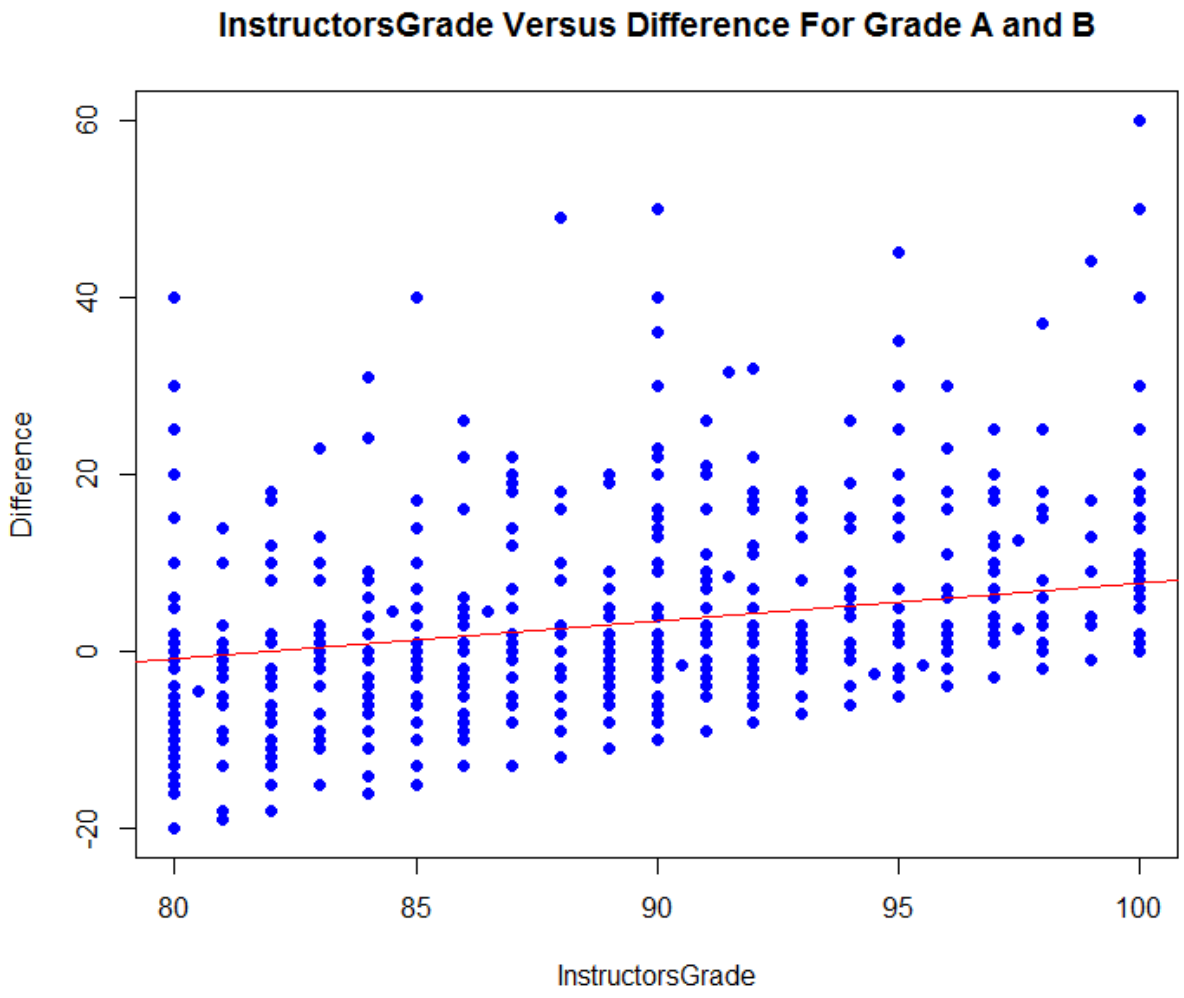
```
Residuals:
   Min     1Q Median     3Q    Max
-18.08  -7.80  -1.10   4.53  45.53

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       -55.861     13.968   -4.00 0.000072 ***
InstructorsGrade    0.674      0.167    4.04 0.000062 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 12 on 535 degrees of freedom
Multiple R-squared:  0.0296,   Adjusted R-squared:  0.0277
F-statistic: 16.3 on 1 and 535 DF,  p-value: 0.0000621
```

Grades A and B

Descriptive Statistics

	StudentsGrade	InstructorsGrade	Difference
nbr.val	1386.00	1386.000	1386.00
nbr.null	0.00	0.000	276.00
nbr.na	0.00	0.000	0.00
min	39.00	80.000	-20.00
max	100.00	100.000	60.00
range	61.00	20.000	80.00
sum	120901.00	126299.000	5398.00
median	90.00	91.000	0.00
mean	87.23	91.125	3.89
SE.mean	0.33	0.188	0.32
CI.mean.0.95	0.64	0.370	0.62
var	147.33	49.177	140.15
std.dev	12.14	7.013	11.84
coef.var	0.14	0.077	3.04

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = -10, df = 1000, p-value = 1
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -4.4  Inf
sample estimates:
mean of the differences
                  -3.9
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = -10, df = 1000, p-value <0.0000000000000002
```

```
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf -3.4
sample estimates:
mean of the differences
                  -3.9
```

Interpretation:

Students significantly under estimated their grades.

Regression Analysis

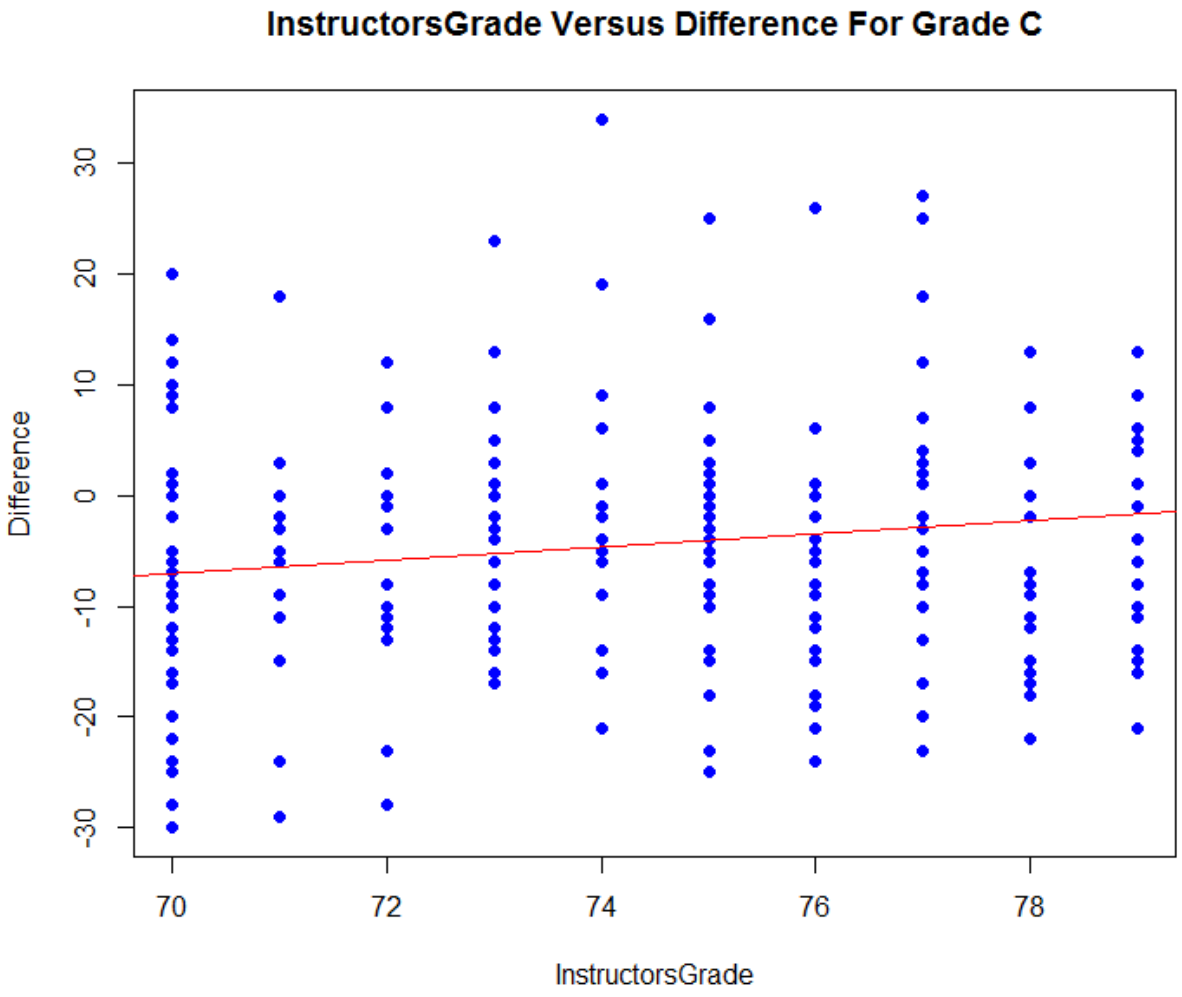
```
Residuals:
   Min     1Q Median     3Q    Max
-19.14  -7.68  -2.68   6.59  52.32

Coefficients:
                 Estimate Std. Error t value            Pr(>|t|)
(Intercept)      -35.0167     4.0124   -8.73 <0.0000000000000002 ***
InstructorsGrade   0.4270     0.0439    9.73 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 12 on 1384 degrees of freedom
Multiple R-squared:  0.064,    Adjusted R-squared:  0.0633
F-statistic: 94.6 on 1 and 1384 DF,  p-value: <0.0000000000000002
```

Grade C

Descriptive Statistics

```
             StudentsGrade InstructorsGrade Difference
nbr.val             407.00          407.000     407.00
nbr.null              0.00            0.000      22.00
nbr.na                0.00            0.000       0.00
min                  40.00           70.000     -30.00
max                 100.00           79.000      34.00
range                60.00            9.000      64.00
sum               31977.00        30090.000   -1887.00
median               80.00           74.000      -5.00
mean                 78.57           73.931      -4.64
SE.mean               0.67            0.160       0.67
CI.mean.0.95          1.31            0.315       1.32
var                 181.51           10.468     183.49
std.dev              13.47            3.235      13.55
coef.var              0.17            0.044      -2.92
```

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 7, df = 400, p-value = 0.00000000001
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 3.5 Inf
sample estimates:
mean of the differences
                    4.6
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 7, df = 400, p-value = 1
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
```

```
 -Inf  5.7
sample estimates:
mean of the differences
                    4.6
```

Interpretation:

Students significantly over estimated their grades.

Regression Analysis

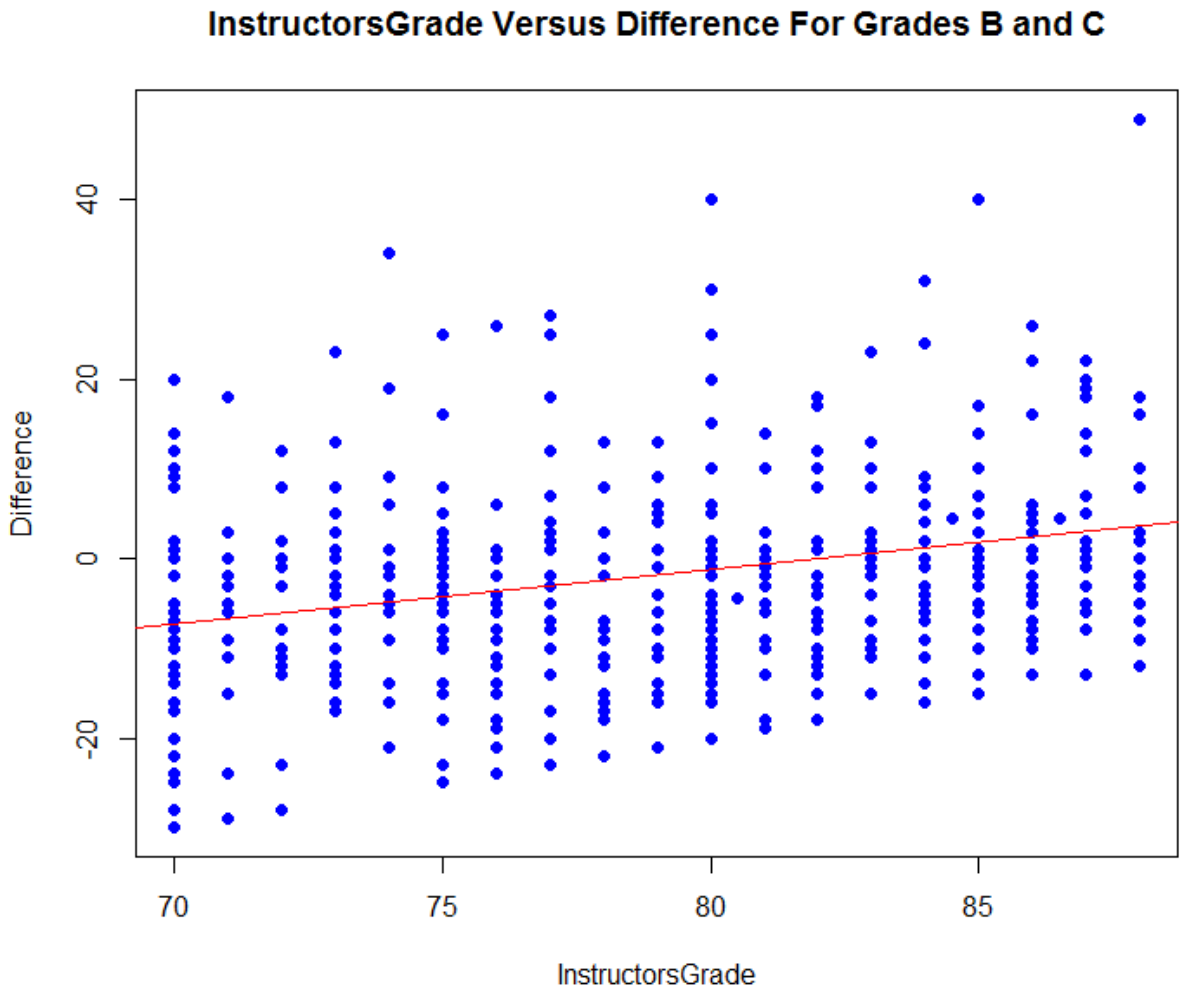
```
Residuals:
   Min     1Q Median     3Q    Max
-23.03  -8.61  -0.81   6.89  38.60

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       -48.597     15.239   -3.19   0.0015 **
InstructorsGrade    0.595      0.206    2.89   0.0041 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13 on 405 degrees of freedom
Multiple R-squared:  0.0202,   Adjusted R-squared:  0.0178
F-statistic: 8.34 on 1 and 405 DF,  p-value: 0.00409
```

Grades B and C

Descriptive Statistics

	StudentsGrade	InstructorsGrade	Difference
nbr.val	901.00	901.000	901.00
nbr.null	0.00	0.000	72.00
nbr.na	0.00	0.000	0.00
min	39.00	70.000	-30.00
max	100.00	88.000	49.00
range	61.00	18.000	79.00
sum	72831.00	71140.500	-1690.50
median	81.00	80.000	-2.00
mean	80.83	78.957	-1.88
SE.mean	0.43	0.182	0.44
CI.mean.0.95	0.84	0.357	0.86
var	166.66	29.782	172.56
std.dev	12.91	5.457	13.14
coef.var	0.16	0.069	-7.00

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 4, df = 900, p-value = 0.00001
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 1.2 Inf
sample estimates:
mean of the differences
                    1.9
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 4, df = 900, p-value = 1
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
```

```
 -Inf  2.6
sample estimates:
mean of the differences
                    1.9
```

Interpretation:

Students significantly over estimated their grades.

Regression Analysis

```
Residuals:
   Min     1Q Median     3Q    Max
-22.76  -8.15  -0.75   6.04  45.46

Coefficients:
                  Estimate Std. Error t value           Pr(>|t|)
(Intercept)      -49.1771     6.1539   -7.99 0.0000000000000041 ***
InstructorsGrade   0.5991     0.0778    7.70 0.0000000000000346 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13 on 899 degrees of freedom
Multiple R-squared:  0.0619,   Adjusted R-squared:  0.0609
F-statistic: 59.4 on 1 and 899 DF,  p-value: 0.0000000000000346
```

Grades Between 75 and 85

Descriptive Statistics

Scatter Plot

T Test

Interpretation:

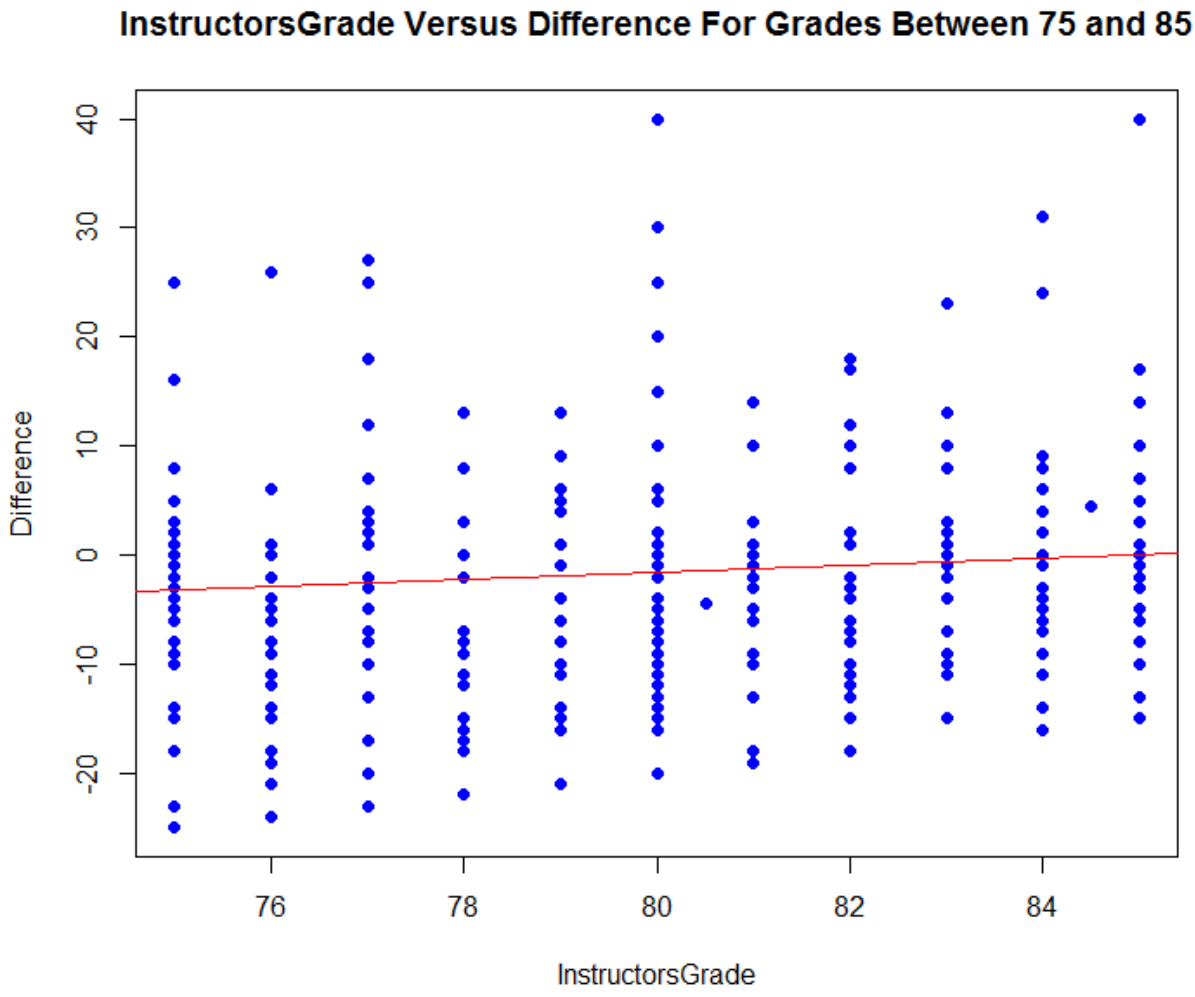
Regression Analysis

Grades Below C

Descriptive Statistics

	StudentsGrade	InstructorsGrade	Difference
nbr.val	568.00	568.000	568.00
nbr.null	0.00	0.000	52.00
nbr.na	0.00	0.000	0.00
min	40.00	75.000	-25.00
max	100.00	85.000	40.00
range	60.00	10.000	65.00
sum	46421.00	45560.000	-861.00
median	82.00	80.000	-2.00
mean	81.73	80.211	-1.52
SE.mean	0.51	0.122	0.51
CI.mean.0.95	1.01	0.240	1.00
var	150.23	8.476	147.36
std.dev	12.26	2.911	12.14
coef.var	0.15	0.036	-8.01

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 3, df = 600, p-value = 0.002
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.68  Inf
sample estimates:
mean of the differences
                    1.5
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 3, df = 600, p-value = 1
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
```

```
 -Inf  2.4
sample estimates:
mean of the differences
                    1.5
```

Interpretation:

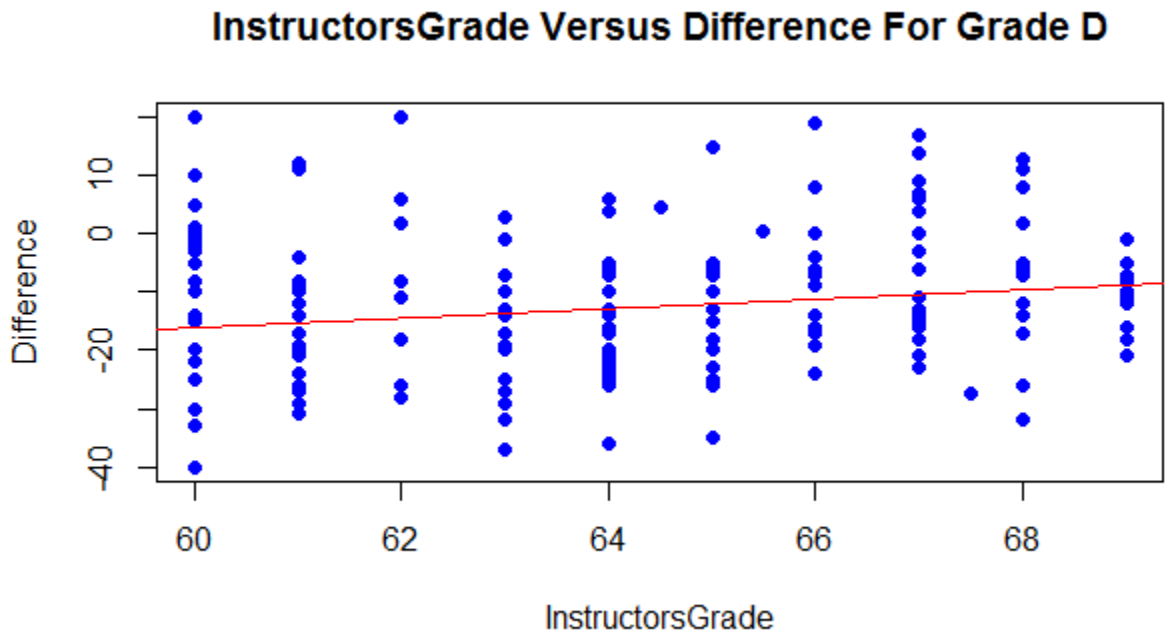
Students significantly over estimated their grades.

Grade D

Descriptive

	StudentsGrade	InstructorsGrade	Difference
nbr.val	310.00	310.000	310.00
nbr.null	0.00	0.000	14.00
nbr.na	0.00	0.000	0.00
min	40.00	60.000	-40.00
max	100.00	69.000	20.00
range	60.00	9.000	60.00
sum	23790.00	19721.500	-4068.50
median	80.00	64.000	-14.00
mean	76.74	63.618	-13.12
SE.mean	0.72	0.168	0.74
CI.mean.0.95	1.43	0.330	1.45
var	162.79	8.736	167.76
std.dev	12.76	2.956	12.95
coef.var	0.17	0.046	-0.99

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 20, df = 300, p-value <0.0000000000000002
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  12 Inf
sample estimates:
mean of the differences
                     13
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 20, df = 300, p-value = 1
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf   14
sample estimates:
```

```
mean of the differences
                    13
```

Interpretation: Students significantly over estimated their grades.

Regression Analysis

```
Residuals:
   Min     1Q Median     3Q    Max
-24.04  -8.18  -2.32   7.18  35.96

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       -63.051     15.645   -4.03  0.00007 ***
InstructorsGrade    0.785      0.246    3.19   0.0015 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13 on 308 degrees of freedom
Multiple R-squared:  0.0321,    Adjusted R-squared:  0.0289
F-statistic: 10.2 on 1 and 308 DF,  p-value: 0.00155
```

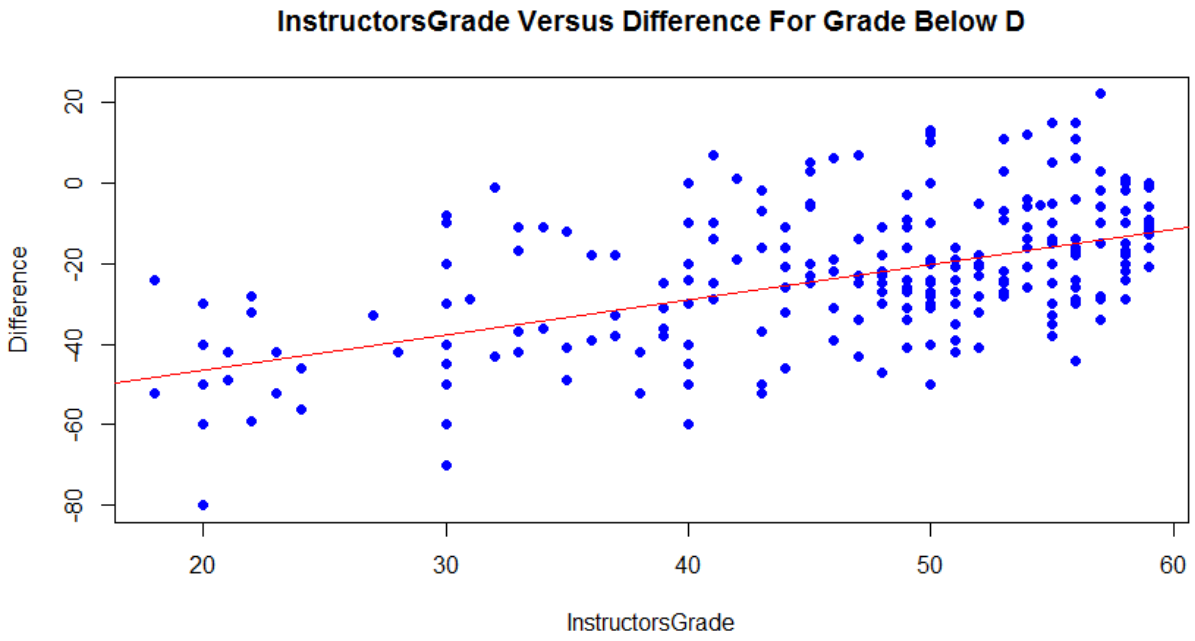
Grades Below D

Descriptive

	StudentsGrade	InstructorsGrade	Difference
nbr.val	464.00	464.00	464.00
nbr.null	0.00	0.00	19.00
nbr.na	0.00	0.00	0.00
min	33.00	18.00	-80.00
max	100.00	59.00	22.00
range	67.00	41.00	102.00
sum	32358.00	21378.50	-10979.50
median	70.00	50.00	-24.00
mean	69.74	46.07	-23.66
SE.mean	0.71	0.45	0.81
CI.mean.0.95	1.39	0.89	1.60
var	233.36	96.04	306.07

std.dev	15.28	9.80	17.49
coef.var	0.22	0.21	-0.74

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 30, df = 500, p-value <0.0000000000000002
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  22 Inf
sample estimates:
mean of the differences
                     24
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 30, df = 500, p-value = 1
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
```

```
 -Inf   25
sample estimates:
mean of the differences
                    24
```

Interpretation: Students significantly over estimated their grades.

Regression Analysis

```
Residuals:
   Min     1Q Median     3Q    Max
-33.43  -9.79  -0.39  10.70  36.06


Coefficients:
                 Estimate Std. Error t value            Pr(>|t|)
(Intercept)      -64.1398     3.4056   -18.8 <0.0000000000000002 ***
InstructorsGrade   0.8785     0.0723    12.2 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 15 on 462 degrees of freedom
Multiple R-squared:  0.242,     Adjusted R-squared:  0.241
F-statistic:  148 on 1 and 462 DF,  p-value: <0.0000000000000002


Analysis By Courses
Only College Algebra and Calculus II are repeated in both the semesters, so
we do combined analysis for these courses.
```

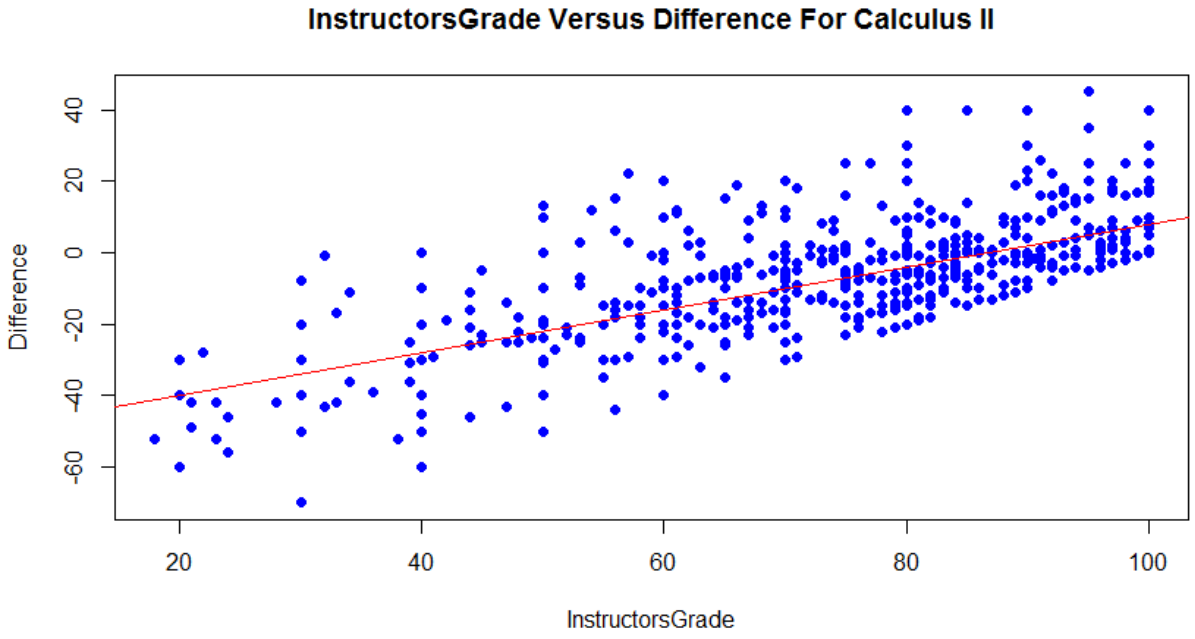
Calculus II

Descriptive

	StudentsGrade	InstructorsGrade	Difference
nbr.val	801.00	801.00	801.00
nbr.null	0.00	0.00	117.00
nbr.na	0.00	0.00	0.00
min	33.00	18.00	-70.00
max	100.00	100.00	45.00
range	67.00	82.00	115.00
sum	66038.00	61138.50	-4899.50
median	85.00	80.00	-3.00

mean	82.44	76.33	-6.12
SE.mean	0.55	0.69	0.63
CI.mean.0.95	1.09	1.35	1.24
var	245.14	380.98	320.34
std.dev	15.66	19.52	17.90
coef.var	0.19	0.26	-2.93

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 10, df = 800, p-value <0.0000000000000002
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 5.1 Inf
sample estimates:
mean of the differences
                    6.1
        Paired t-test
```

```
data:  StudentsGrade and InstructorsGrade
t = 10, df = 800, p-value = 1
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf  7.2
sample estimates:
mean of the differences
                    6.1
```

Interpretation: Students significantly over estimated their grades

Regression Analysis

```
Residuals:
   Min     1Q Median     3Q    Max
-36.15  -8.09  -2.09   7.89  43.92

Coefficients:
                 Estimate Std. Error t value            Pr(>|t|)
(Intercept)      -51.8136     1.9358   -26.8 <0.0000000000000002 ***
InstructorsGrade   0.5987     0.0246    24.4 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14 on 799 degrees of freedom
Multiple R-squared:  0.426,     Adjusted R-squared:  0.426
F-statistic:  594 on 1 and 799 DF,  p-value: <0.0000000000000002
```

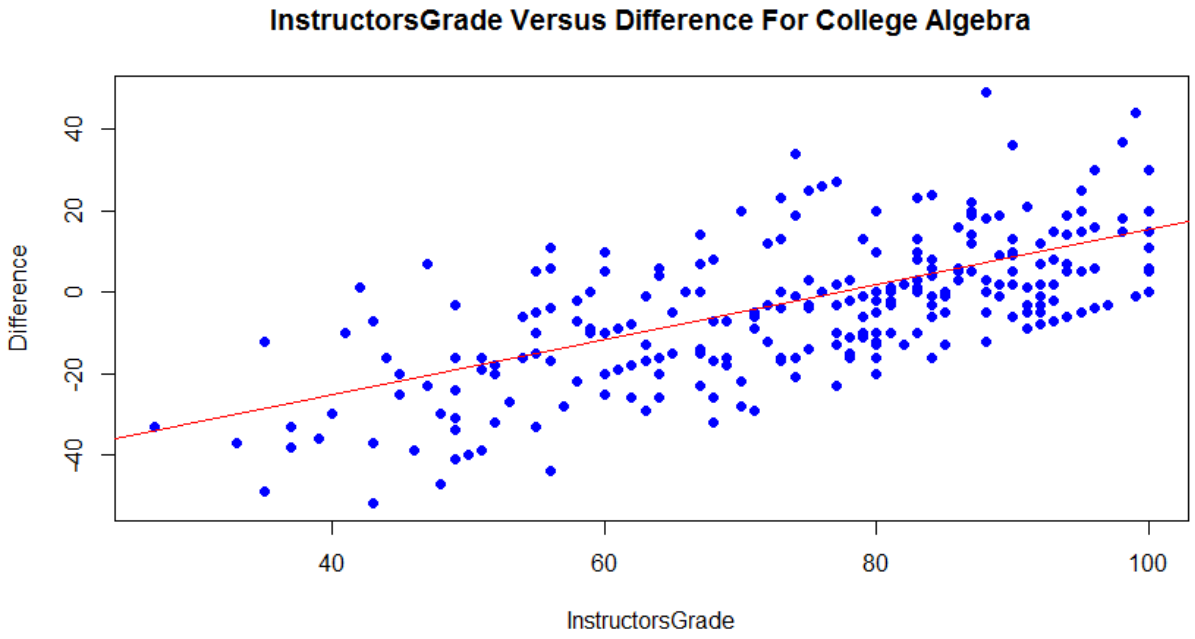
College Algebra

Descriptive

	StudentsGrade	InstructorsGrade	Difference
nbr.val	559.00	559.00	559.00
nbr.null	0.00	0.00	28.00
nbr.na	0.00	0.00	0.00
min	39.00	27.00	-52.00
max	100.00	100.00	49.00
range	61.00	73.00	101.00

sum	42695.00	41676.00	-1019.00
median	80.00	79.00	-1.00
mean	76.38	74.55	-1.82
SE.mean	0.60	0.72	0.73
CI.mean.0.95	1.17	1.42	1.44
var	199.01	293.67	300.83
std.dev	14.11	17.14	17.34
coef.var	0.18	0.23	-9.51

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 2, df = 600, p-value = 0.007
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.61  Inf
sample estimates:
mean of the differences
                    1.8
```

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 2, df = 600, p-value = 1
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf    3
sample estimates:
mean of the differences
                    1.8
```

Interpretation: Students significantly over estimated their grades.

Regression Analysis

```
Residuals:
   Min     1Q Median     3Q    Max
-29.68  -8.46  -1.32   7.58  41.77

Coefficients:
                 Estimate Std. Error t value            Pr(>|t|)
(Intercept)       -52.025      2.449   -21.2 <0.0000000000000002 ***
InstructorsGrade    0.673      0.032    21.0 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13 on 557 degrees of freedom
Multiple R-squared:  0.443,     Adjusted R-squared:  0.442
F-statistic:  442 on 1 and 557 DF,  p-value: <0.0000000000000002
```

Analysis by Gender

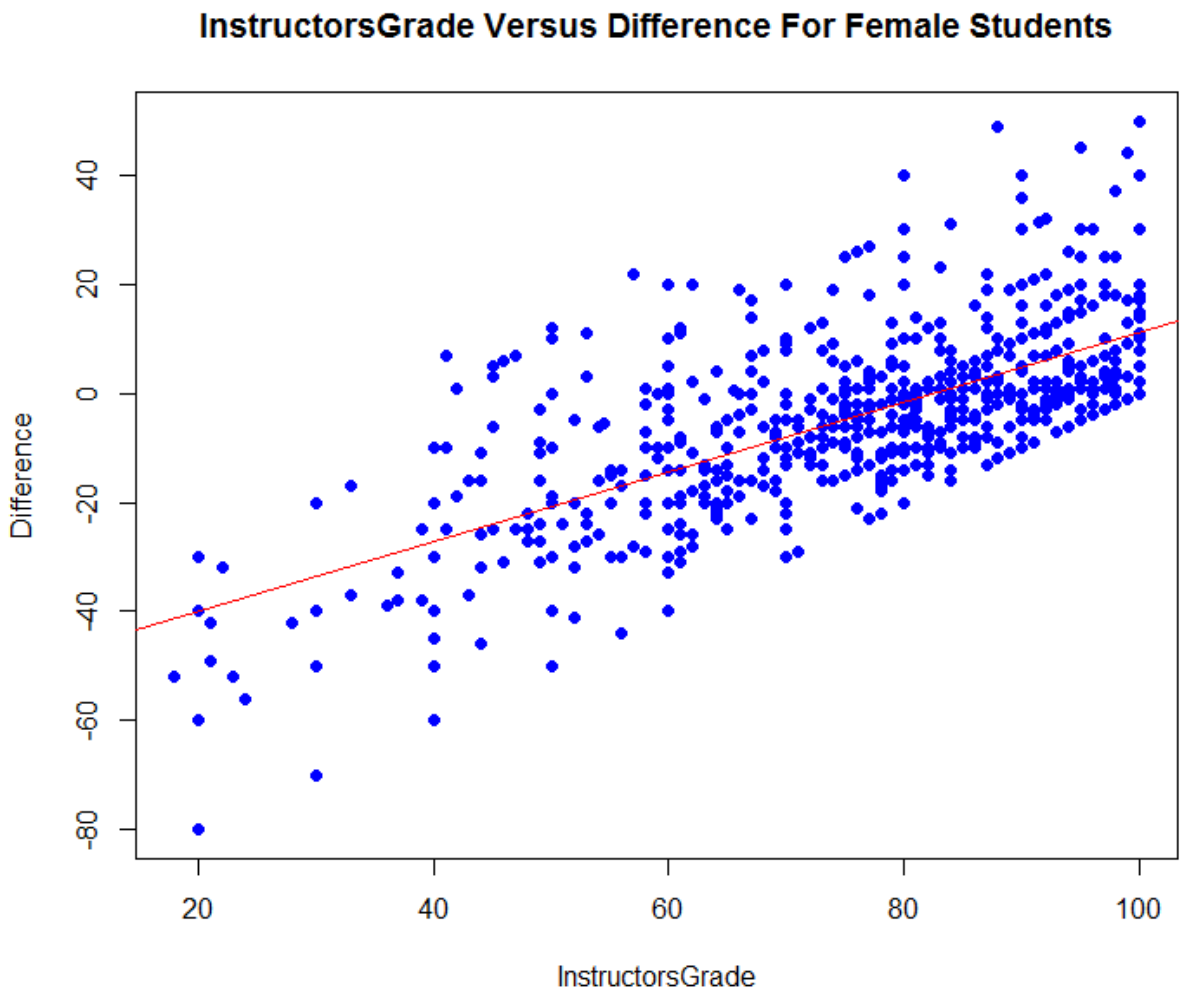
Female Students

Descriptive Statistics

	StudentsGrade	InstructorsGrade	Difference
nbr.val	1233.00	1233.00	1233.00
nbr.null	0.00	0.00	152.00
nbr.na	0.00	0.00	0.00
min	34.00	18.00	-80.00
max	100.00	100.00	50.00

range	66.00	82.00	130.00
sum	99876.00	96388.50	-3487.50
median	81.00	80.00	0.00
mean	81.00	78.17	-2.83
SE.mean	0.41	0.51	0.49
CI.mean.0.95	0.81	1.00	0.96
var	209.04	318.99	297.68
std.dev	14.46	17.86	17.25
coef.var	0.18	0.23	-6.10

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
```

```
t = 6, df = 1000, p-value = 0.000000005
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
   2 Inf
sample estimates:
mean of the differences
                    2.8
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 6, df = 1000, p-value = 1
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf  3.6
sample estimates:
mean of the differences
                    2.8
```

Interpretation:

Students significantly over estimated their grades

Regression Analysis

```
Residuals:
   Min     1Q Median     3Q    Max
-40.00  -9.01  -1.78   6.72  45.55

Coefficients:
                 Estimate Std. Error t value            Pr(>|t|)
(Intercept)      -52.7769     1.6559   -31.9 <0.0000000000000002 ***
InstructorsGrade   0.6389     0.0207    30.9 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13 on 1231 degrees of freedom
Multiple R-squared:  0.437,    Adjusted R-squared:  0.437
F-statistic:  957 on 1 and 1231 DF,  p-value: <0.0000000000000002
```

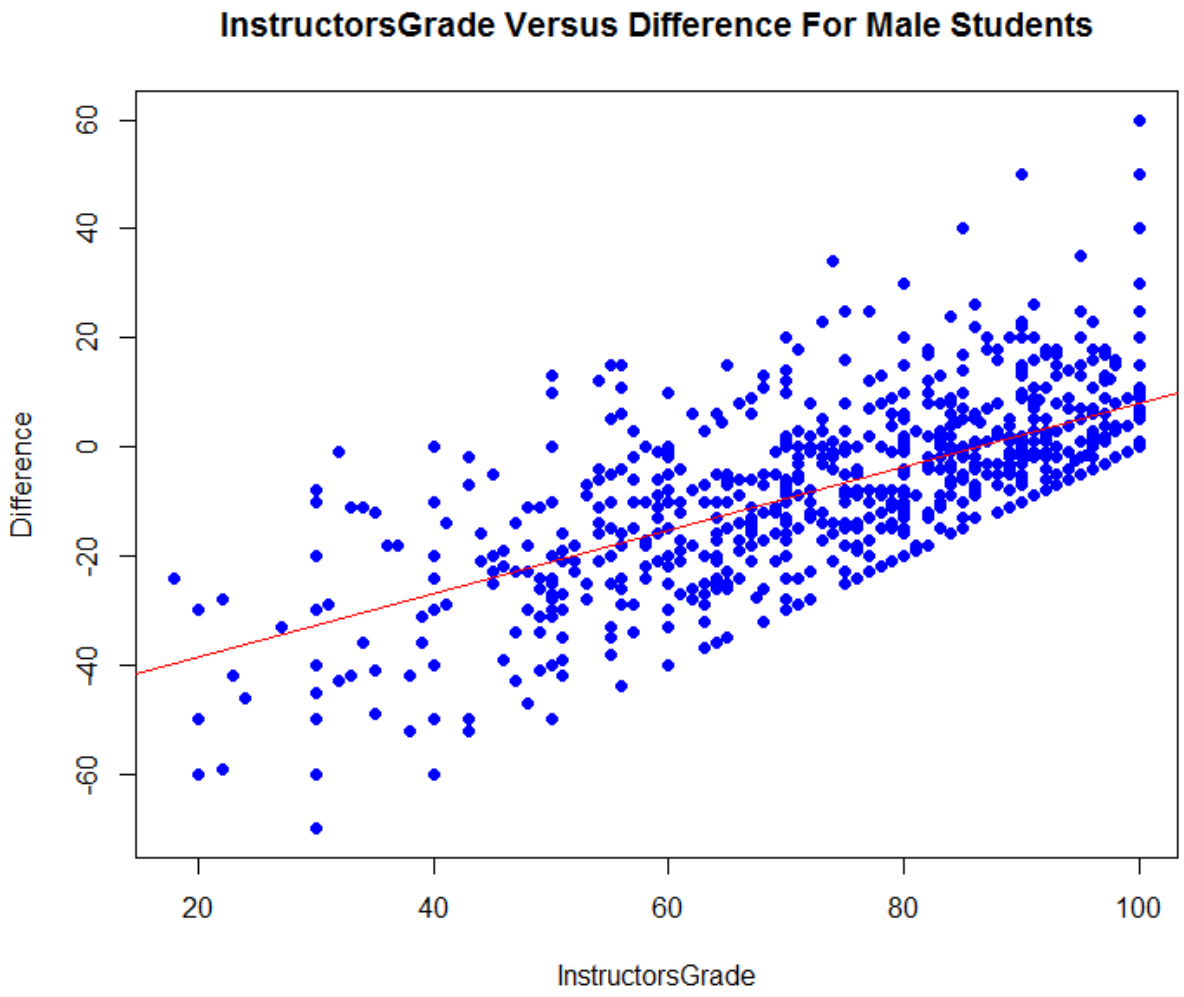
Male Students

Descriptive Statistics

```
             StudentsGrade InstructorsGrade Difference
nbr.val            1334.00          1334.00    1334.00
nbr.null              0.00             0.00     179.00
nbr.na                0.00             0.00       0.00
min                  33.00            18.00     -70.00
max                 100.00           100.00      60.00
range                67.00            82.00     130.00
sum              109150.00        101100.50   -8049.50
median               82.50            80.00      -4.00
mean                 81.82            75.79      -6.03
SE.mean               0.41             0.53       0.46
```

CI.mean.0.95	0.80	1.04	0.91
var	223.81	376.98	285.04
std.dev	14.96	19.42	16.88
coef.var	0.18	0.26	-2.80

Scatter Plot



T Test

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 10, df = 1000, p-value <0.0000000000000002
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 5.3 Inf
sample estimates:
mean of the differences
                      6
```

```
        Paired t-test

data:  StudentsGrade and InstructorsGrade
t = 10, df = 1000, p-value = 1
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf  6.8
sample estimates:
mean of the differences
                      6
```

Interpretation:

Students significantly over estimated their grades

Regression Analysis

```
Residuals:
   Min     1Q Median     3Q    Max
-37.35  -8.04  -1.54   7.70  51.96

Coefficients:
                 Estimate Std. Error t value            Pr(>|t|)
(Intercept)      -50.0819     1.3864   -36.1 <0.0000000000000002 ***
InstructorsGrade   0.5812     0.0177    32.8 <0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13 on 1332 degrees of freedom
Multiple R-squared:  0.447,    Adjusted R-squared:  0.446
F-statistic: 1.08e+03 on 1 and 1332 DF,  p-value: <0.0000000000000002

Logistic Regression Analysis
Dependent Variable=Behavior(overestimate=1, underestimate=0)-Binomial Variabl
e
Independent variable/predictor variables:
Semester (Spring=1, Fall=2)-Categorical
CourseLevel(lower level courses=1, upper level courses=2)-Categorical
Gender (Female=1, Male=2)-Categorical
InstructorsGrade (Continuous Variable)-Continuous
```

```
Analysis Result:
Deviance Residuals:
   Min      1Q  Median      3Q     Max
-2.966  -0.861   0.261   0.868   1.917

Coefficients:
             Estimate Std. Error z value             Pr(>|z|)
(Intercept)   8.60209    0.52607   16.35 < 0.0000000000000002 ***
Gender1       0.14224    0.10421    1.36                 0.17
Semester2     0.81653    0.17775    4.59            0.0000044 ***
Grade        -0.11339    0.00736  -15.40 < 0.0000000000000002 ***
CourseLevel2  0.43345    0.10818    4.01            0.0000616 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2998.5  on 2235  degrees of freedom
Residual deviance: 2243.6  on 2231  degrees of freedom
AIC: 2254

Number of Fisher Scoring iterations: 5
```

Interpretation: When the gender changes from female to male, the log(probability of under estimate) increases by 0.14, but this is not significant as the p value is more than 0.05. This means, the gender analysis is not suitable for the logistic regression. When the grade increases by 1, the log(probability of under estimate) decreases by 0.11. When the course level increases by 1, the log(probability of under estimate) increases by 0.4.

Correlation Coefficient of the combined data:

```
     Pearson's product-moment correlation


data:  x and y

t = 30, df = 3000, p-value <0.0000000000000002

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

 0.47 0.53

sample estimates:

cor

0.5
```