भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Final Exam Report

Submitted by

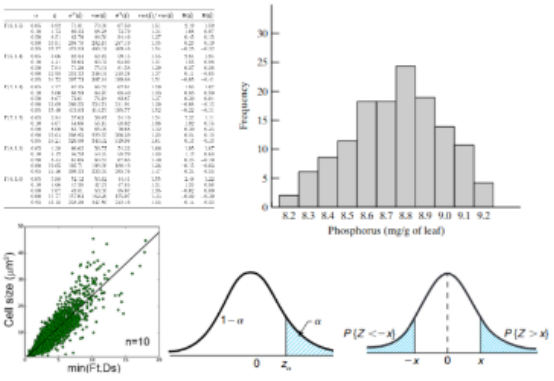
Gaurav Kumar Singh

Towards partial fulfillment of the requirements for the subject

Statistical Inference in Bioengineering

In the department of

Biomedical engineering



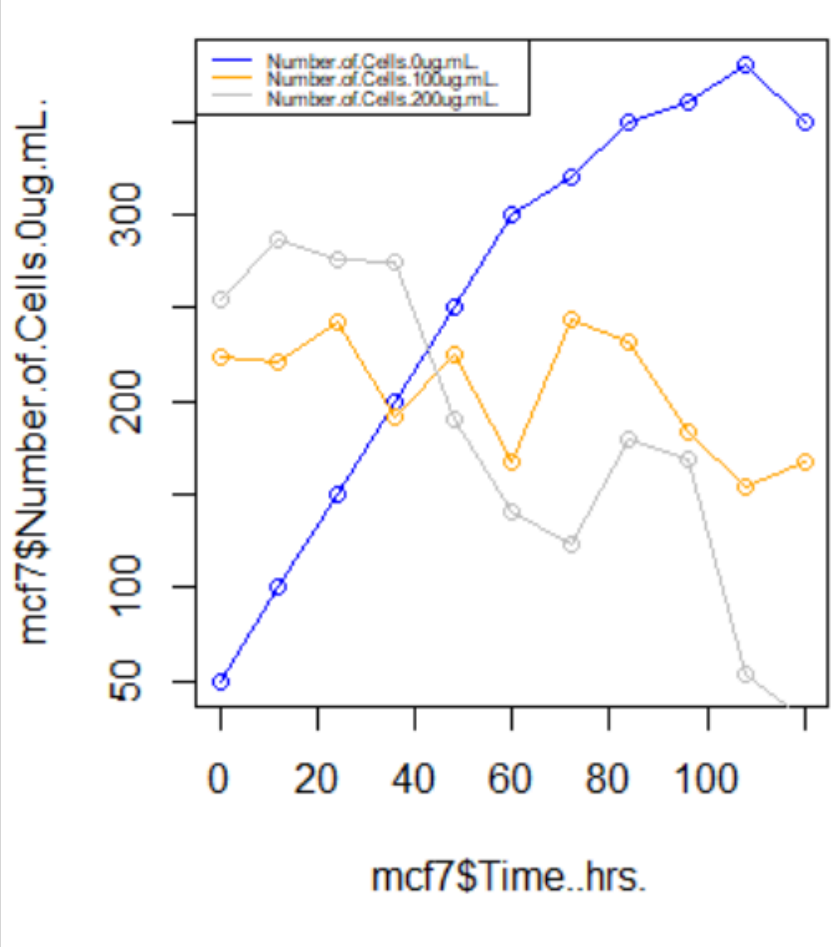
Question 1.You are expected to explore the data and identify patterns, if any.

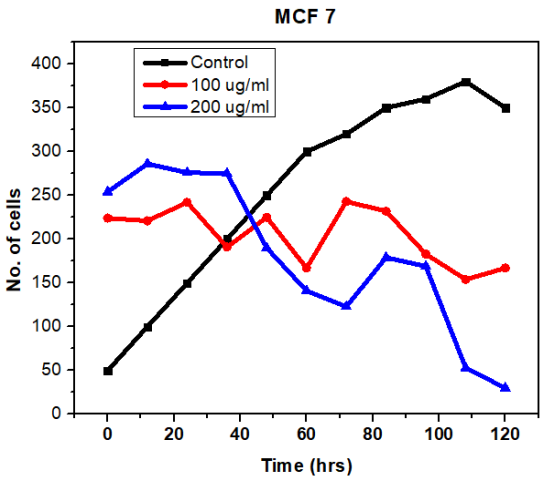
Answer 1. From the given data we can infer that anti-cancer drug D1344320 with different concentrations shows different results in every cell line.

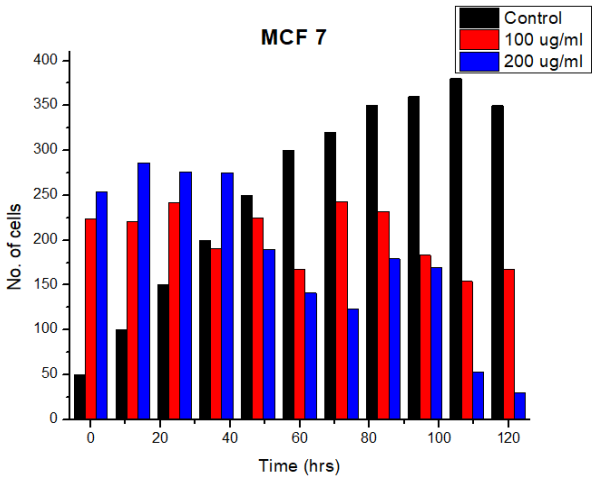
From the plotted graph we can infer that when the *concentration of drug is 0* in MCF7 cell line then the cell line ***increases exponentially,***which shows that it had no effects on the cell line.

When the concentration was changed to 100 we can observe that there is ***slight decrease*** in the cell line, which means the drug is working at this concentration but not effectively.

Now, when the concentration is changed to 200 we can observe that there is a high decrease in the cell line, hence, we can infer that the ***drug worked on MCF7*** cell line at ***drug concentration of 200.***



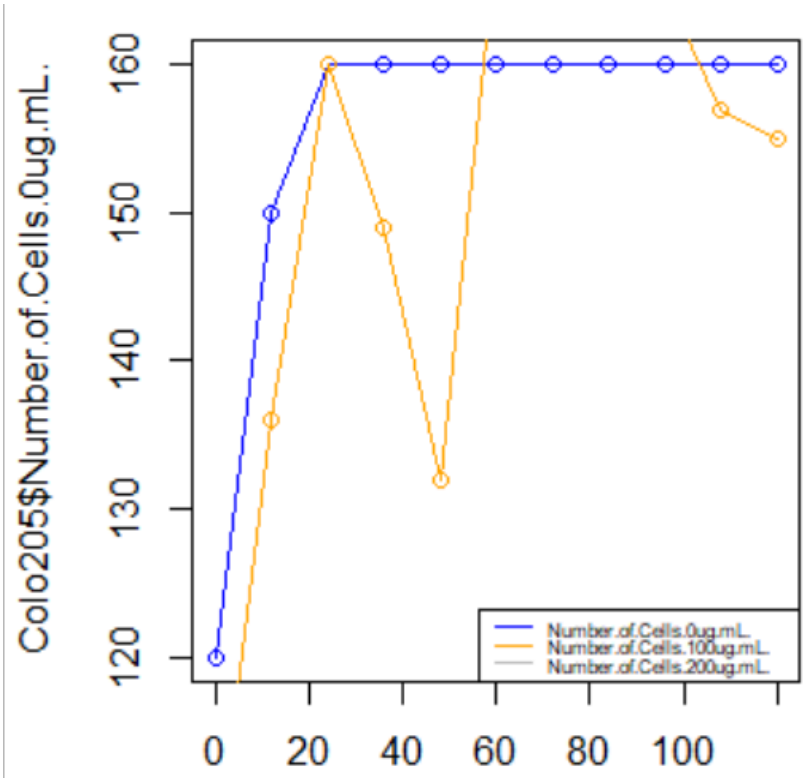


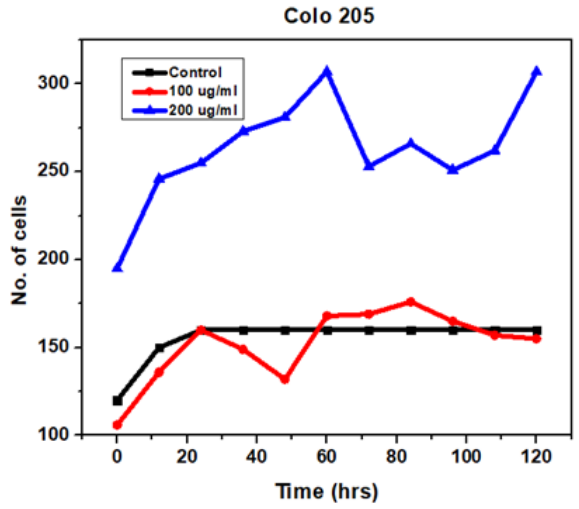


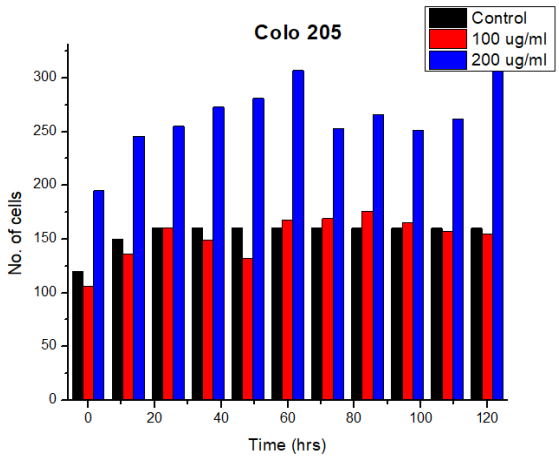
Now, coming to the second cell line i.e. ***Colo205*** when the anti cancer drug is provided to this cell line at *drug concentration 0* then we can observe that there is ***explicitly no change*** in cell line. The drug has no effect at this concentration on this cell line.

When the concentration is *increased to 100* then we can observe that there is slight increase in number of cells and ***then slight decrease with constant cell line***. Hence we can infer that this drug does not work on this particular cell line.

Now, when the concentration is *increased to 200* then we can observe that the cell line is ***increased*** at the end of result, which shows that the drug affects this cell line but ***not effectively because it does not decrease the cell line.***





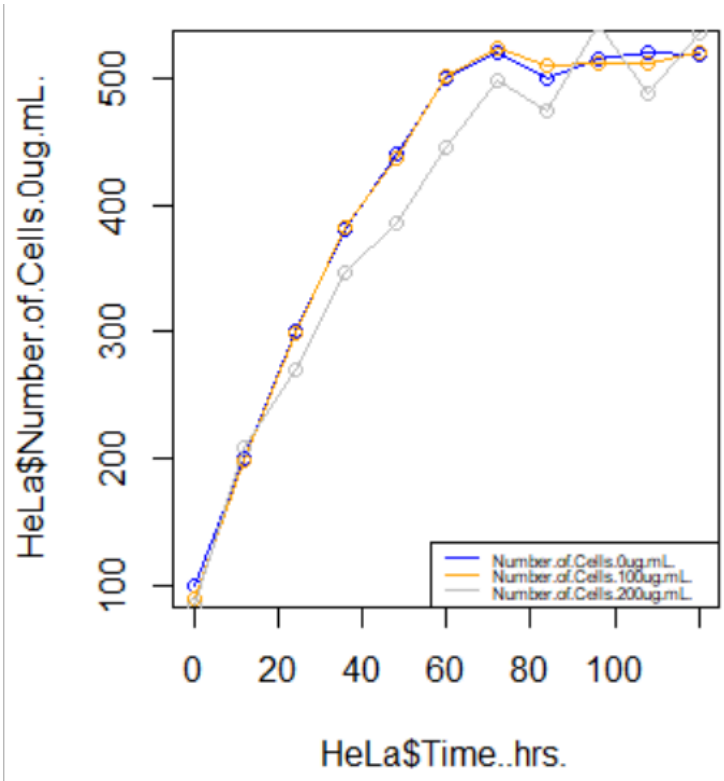


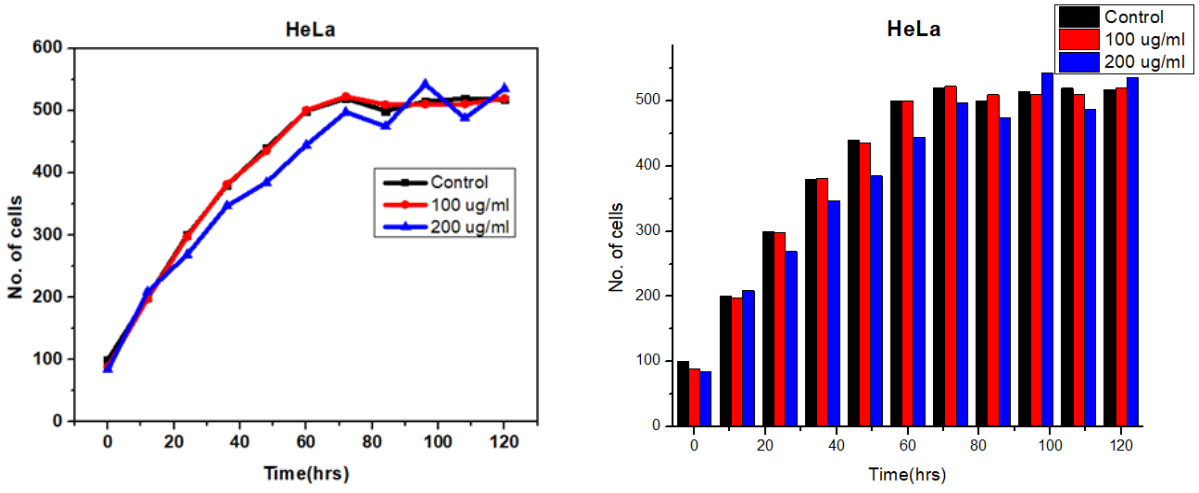
Now coming to the third cell line i.e. ***HeLa*** when anti cancer drug was given to this cell line at drug *concentration level 0* then we can observe that the cell line increases which shows that the ***drug has no effect*** at all on this cell line at 0 concentration level.

When the concentration is *increased to 100* then we can observe that there is an increase in the number of cells, ***hence no effect at this concentration*** too.

Now, when the concentration is *increased to 200* than we can infer that there is ***again increase in the cell line***

From the above data we can infer that this ***anti drug has not worked on this particular cell line.***

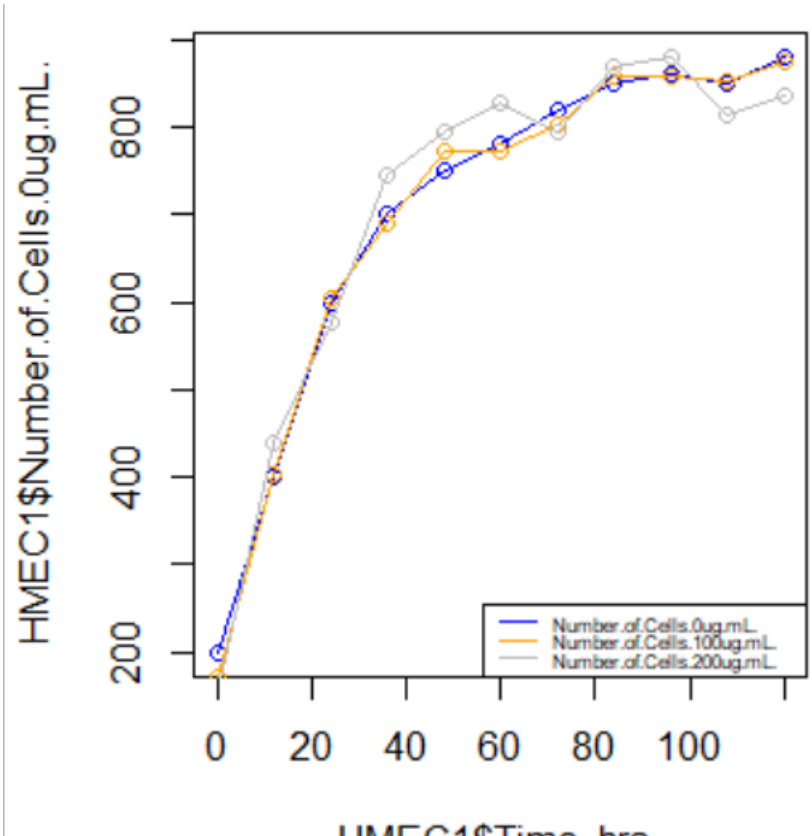


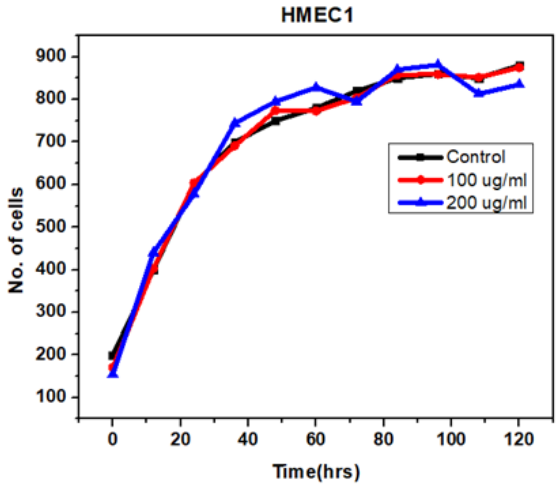


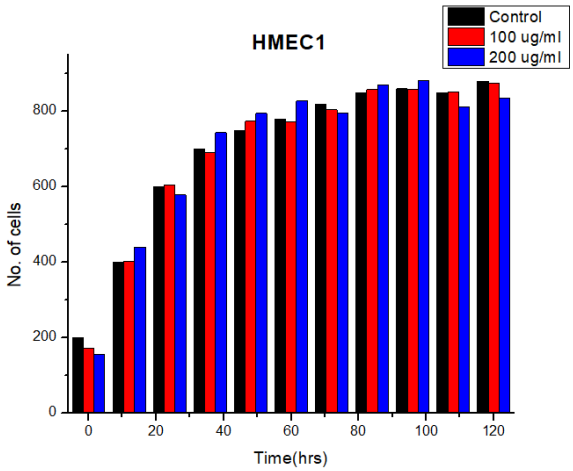
Now coming to the fourth cell line that is ***HMEC-1,*** when an anti cancer cell is given to this cell line at *drug concentration level 0* then we can observe that there is an increase in the number of cells which infers that ***drug has no effect*** on this cell line.

When the concentration is *increased to 100* then we can observe an increase in the number of cells, which gives us the results that the ***drug is not working at this concentration too.***

When the concentration is *increased to 200* then we can observe that there is an increase in the number of cells, from this we can infer that this ***drug shows no effect*** on this cell type at any drug concentration.



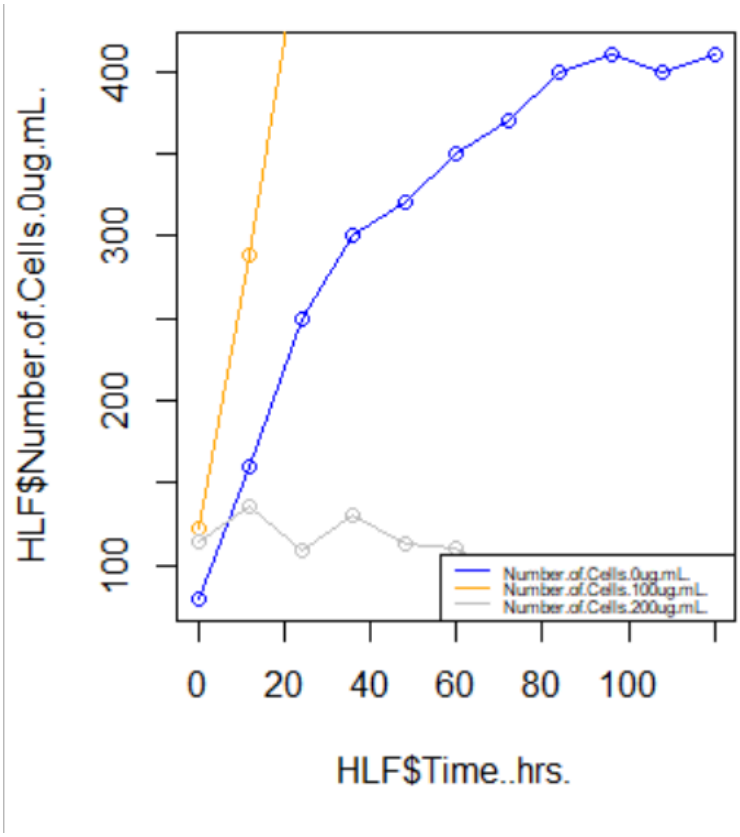


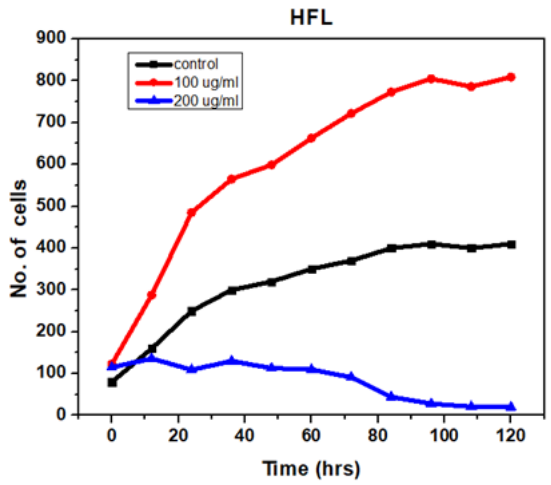


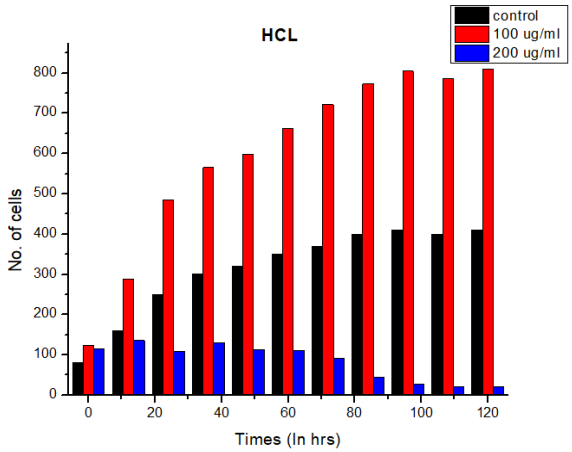
Now coming to the fifth and last cell line that is ***HLF*** when *drug concentration at level 0* was given to this cell line than we can observe that there is a ***slight increase in the cell line.***

Now increasing the *concentration of the drug to 100* we can observe that the cell line increases explicitly. Hence the ***drug does not work at this concentration.***

When the concentration is *increased to 200* than we can observe that the number of cells or ***cell lines decrease*** , hence the ***drug works on this cell*** line at this concentration.







From all the above data and graphs we can summarize that

The given data set shows the effect of a drug on cancer cells of a particular type. The treated drug shows effectiveness on MCF7 at 200 ug/ml and HFL at 200 ug/ml. This drug is not suitable to treat HeLa and HMEC1 cancer cells. Whereas, In the case of Colo 205 drug at 200 ug/ml induce the growth of cancer cells.

Question 2: Perform appropriate statistical tests to infer the effectiveness of the drug on different cells.

Answer 2: If we infer the data set which is provided, we can summarize that the control parameter is drug concentration i.e. qualitative data, hence we can apply regression on this data set to get our answer.

(I) For MCF7

```
> #linear regression
> covariance <- cov(mcf7)
> correlation <- cor(mcf7)
> regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.
+                  + Number.of.Cells.100ug.mL.
+                  + Number.of.Cells.200ug.mL., data = mcf7)
> covariance
                         Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                    1584.0                4344.000                    -756.000
Number.of.Cells.0ug.mL.       4344.0               13107.273                   -1855.727
Number.of.Cells.100ug.mL.     -756.0               -1855.727                    1072.473
Number.of.Cells.200ug.mL.    -3158.4               -8378.818                    1722.982
                         Number.of.Cells.200ug.mL.
Time..hrs.                             -3158.400
Number.of.Cells.0ug.mL.                -8378.818
Number.of.Cells.100ug.mL.               1722.982
Number.of.Cells.200ug.mL.               7857.255
> correlation
                         Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                1.0000000                0.9533581                -0.5800308
Number.of.Cells.0ug.mL.   0.9533581                1.0000000                -0.4949540
Number.of.Cells.100ug.mL. -0.5800308               -0.4949540                 1.0000000
Number.of.Cells.200ug.mL. -0.8952702               -0.8256409                 0.5935430
                         Number.of.Cells.200ug.mL.
Time..hrs.                            -0.8952702
Number.of.Cells.0ug.mL.               -0.8256409
Number.of.Cells.100ug.mL.              0.5935430
Number.of.Cells.200ug.mL.              1.0000000
```

```
> regression

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = mcf7)

Coefficients:
               (Intercept)    Number.of.Cells.0ug.mL.  Number.of.Cells.100ug.mL.
                   41.8616                    0.2335                    -0.0851
Number.of.Cells.200ug.mL.
                   -0.1343
```

```
> summary(regression)

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = mcf7)

Residuals:
     Min       1Q   Median       3Q      Max
-18.7783  -4.8907   0.6102   4.4703  14.6408

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               41.86161   32.83722   1.275  0.24305
Number.of.Cells.0ug.mL.    0.23354    0.05270   4.432  0.00304 **
Number.of.Cells.100ug.mL. -0.08510    0.12915  -0.659  0.53099
Number.of.Cells.200ug.mL. -0.13427    0.07348  -1.827  0.11040
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.76 on 7 degrees of freedom
Multiple R-squared:  0.9488,    Adjusted R-squared:  0.9269
F-statistic: 43.24 on 3 and 7 DF,  p-value: 6.926e-05
```

(II) For Colo205

```
> #linear regression
> covariance <- cov(Colo205)
> correlation <- cor(Colo205)
> regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.
+                  + Number.of.Cells.100ug.mL.
+                  + Number.of.Cells.200ug.mL., data = Colo205)
> covariance
                         Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                     1584                288.0000                  522.0000
Number.of.Cells.0ug.mL.         288                147.2727                  200.4545
Number.of.Cells.100ug.mL.       522                200.4545                  416.8909
Number.of.Cells.200ug.mL.       684                290.3636                  352.2727
                         Number.of.Cells.200ug.mL.
Time..hrs.                               684.0000
Number.of.Cells.0ug.mL.                  290.3636
Number.of.Cells.100ug.mL.                352.2727
Number.of.Cells.200ug.mL.                952.6182
```

```
> correlation
                         Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                1.0000000              0.5962848                 0.6423648
Number.of.Cells.0ug.mL.   0.5962848              1.0000000                 0.8089907
Number.of.Cells.100ug.mL. 0.6423648              0.8089907                 1.0000000
Number.of.Cells.200ug.mL. 0.5568255              0.7752140                 0.5589956
                         Number.of.Cells.200ug.mL.
Time..hrs.                               0.5568255
Number.of.Cells.0ug.mL.                  0.7752140
Number.of.Cells.100ug.mL.                0.5589956
Number.of.Cells.200ug.mL.                1.0000000
> regression

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = Colo205)

Coefficients:
              (Intercept)    Number.of.Cells.0ug.mL.  Number.of.Cells.100ug.mL.
                -164.5915                   -0.2755                     1.0282
Number.of.Cells.200ug.mL.
                   0.4218



> summary(regression)

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = Colo205)

Residuals:
    Min      1Q  Median      3Q     Max
-39.390 -24.673   0.199  17.787  44.742

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -164.5915   154.6282  -1.064    0.322
Number.of.Cells.0ug.mL.     -0.2755     2.0482  -0.135    0.897
Number.of.Cells.100ug.mL.    1.0282     0.9274   1.109    0.304
Number.of.Cells.200ug.mL.    0.4218     0.5709   0.739    0.484

Residual standard error: 34.6 on 7 degrees of freedom
Multiple R-squared:  0.4709,    Adjusted R-squared:  0.2441
F-statistic: 2.076 on 3 and 7 DF,  p-value: 0.1919
```

(III) For HeLa

```
> #linear regression
> covariance <- cov(HeLa)
> correlation <- cor(HeLa)
> regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.
+                  + Number.of.Cells.100ug.mL.
+                  + Number.of.Cells.200ug.mL., data = HeLa)
> covariance
                         Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                    1584.0                 5202.00                    5266.80
Number.of.Cells.0ug.mL.       5202.0                21716.27                   22113.71
Number.of.Cells.100ug.mL.     5266.8                22113.71                   22545.36
Number.of.Cells.200ug.mL.     5480.4                21439.60                   21831.30
                         Number.of.Cells.200ug.mL.
Time..hrs.                                5480.4
Number.of.Cells.0ug.mL.                  21439.6
Number.of.Cells.100ug.mL.                21831.3
Number.of.Cells.200ug.mL.                21914.4
> correlation
                         Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                1.0000000               0.8869520                  0.8813342
Number.of.Cells.0ug.mL.   0.8869520               1.0000000                  0.9994023
Number.of.Cells.100ug.mL. 0.8813342               0.9994023                  1.0000000
Number.of.Cells.200ug.mL. 0.9301862               0.9827866                  0.9821689
                         Number.of.Cells.200ug.mL.
Time..hrs.                               0.9301862
Number.of.Cells.0ug.mL.                  0.9827866
Number.of.Cells.100ug.mL.                0.9821689
Number.of.Cells.200ug.mL.                1.0000000


> regression

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = HeLa)

Coefficients:
              (Intercept)    Number.of.Cells.0ug.mL.  Number.of.Cells.100ug.mL.
                -41.0605                   0.9229                    -1.1169
Number.of.Cells.200ug.mL.
                  0.4598
```

```
> summary(regression)

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = HeLa)

Residuals:
    Min      1Q  Median      3Q     Max
-17.191  -6.820  -5.454  12.176  17.311

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               -41.0605    15.5995  -2.632   0.0338 *
Number.of.Cells.0ug.mL.     0.9229     0.9105   1.014   0.3445
Number.of.Cells.100ug.mL.  -1.1169     0.8782  -1.272   0.2440
Number.of.Cells.200ug.mL.   0.4598     0.1667   2.759   0.0281 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.41 on 7 degrees of freedom
Multiple R-squared:  0.9082,    Adjusted R-squared:  0.8688
F-statistic: 23.08 on 3 and 7 DF,  p-value: 0.0005263
```

(IV) For HMEC1

```
> #linear regression
> covariance <- cov(HMEC1)
> correlation <- cor(HMEC1)
> regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.
+                  + Number.of.Cells.100ug.mL.
+                  + Number.of.Cells.200ug.mL., data = HMEC1)
> covariance
                        Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                  1584.0                7620.00                    7734.00
Number.of.Cells.0ug.mL.     7620.0               47629.09                   48862.73
Number.of.Cells.100ug.mL.   7734.0               48862.73                   50282.42
Number.of.Cells.200ug.mL.   7263.6               48483.09                   49964.63
                        Number.of.Cells.200ug.mL.
Time..hrs.                                7263.60
Number.of.Cells.0ug.mL.                  48483.09
Number.of.Cells.100ug.mL.                49964.63
Number.of.Cells.200ug.mL.                50842.89
```

```
> correlation
                        Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                 1.0000000              0.8772861                 0.8665996
Number.of.Cells.0ug.mL.    0.8772861              1.0000000                 0.9984665
Number.of.Cells.100ug.mL.  0.8665996              0.9984665                 1.0000000
Number.of.Cells.200ug.mL.  0.8093926              0.9852332                 0.9881877
                        Number.of.Cells.200ug.mL.
Time..hrs.                              0.8093926
Number.of.Cells.0ug.mL.                 0.9852332
Number.of.Cells.100ug.mL.               0.9881877
Number.of.Cells.200ug.mL.               1.0000000
> regression

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = HMEC1)

Coefficients:
              (Intercept)    Number.of.Cells.0ug.mL.  Number.of.Cells.100ug.mL.
                -56.0565                   0.5598                    -0.0724
Number.of.Cells.200ug.mL.
                 -0.3198


> summary(regression)

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = HMEC1)

Residuals:
    Min      1Q  Median      3Q     Max
-27.247  -8.676   4.576  11.865  14.568

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               -56.0565    20.2637  -2.766   0.0278 *
Number.of.Cells.0ug.mL.     0.5598     0.4503   1.243   0.2538
Number.of.Cells.100ug.mL.  -0.0724     0.4896  -0.148   0.8866
Number.of.Cells.200ug.mL.  -0.3198     0.1574  -2.031   0.0817 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.95 on 7 degrees of freedom
Multiple R-squared:  0.873,     Adjusted R-squared:  0.8186
F-statistic: 16.04 on 3 and 7 DF,  p-value: 0.001614
```

(V) For HLF

```
> #linear regression
> covariance <- cov(HLF)
> correlation <- cor(HLF)
> regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.
+                  + Number.of.Cells.100ug.mL.
+                  + Number.of.Cells.200ug.mL., data = HLF)
> covariance
                         Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                    1584.0                4008.000                   8305.200
Number.of.Cells.0ug.mL.       4008.0               12045.455                  24723.455
Number.of.Cells.100ug.mL.     8305.2               24723.455                  50837.855
Number.of.Cells.200ug.mL.    -1646.4               -3638.455                  -7651.555
                         Number.of.Cells.200ug.mL.
Time..hrs.                             -1646.400
Number.of.Cells.0ug.mL.                -3638.455
Number.of.Cells.100ug.mL.              -7651.555
Number.of.Cells.200ug.mL.               2076.855
> correlation
                         Time..hrs. Number.of.Cells.0ug.mL. Number.of.Cells.100ug.mL.
Time..hrs.                1.0000000               0.9175686                  0.9255053
Number.of.Cells.0ug.mL.   0.9175686               1.0000000                  0.9990885
Number.of.Cells.100ug.mL. 0.9255053               0.9990885                  1.0000000
Number.of.Cells.200ug.mL. -0.9077259              -0.7274486                 -0.7446514
                         Number.of.Cells.200ug.mL.
Time..hrs.                            -0.9077259
Number.of.Cells.0ug.mL.               -0.7274486
Number.of.Cells.100ug.mL.             -0.7446514
Number.of.Cells.200ug.mL.              1.0000000
```

```
> regression

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = HLF)

Coefficients:
               (Intercept)    Number.of.Cells.0ug.mL.  Number.of.Cells.100ug.mL.
                  34.54791                  0.30381                   -0.05295
Number.of.Cells.200ug.mL.
                  -0.45557
```

```
> summary(regression)

Call:
lm(formula = Time..hrs. ~ Number.of.Cells.0ug.mL. + Number.of.Cells.100ug.mL. +
    Number.of.Cells.200ug.mL., data = HLF)

Residuals:
     Min       1Q   Median       3Q      Max
-11.1634  -4.1509   0.0501   4.5322  12.8372

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               34.54791   18.51194   1.866  0.10425
Number.of.Cells.0ug.mL.    0.30381    0.74134   0.410  0.69419
Number.of.Cells.100ug.mL. -0.05295    0.37097  -0.143  0.89052
Number.of.Cells.200ug.mL. -0.45557    0.11418  -3.990  0.00526 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.947 on 7 degrees of freedom
Multiple R-squared:  0.9646,    Adjusted R-squared:  0.9495
F-statistic: 63.62 on 3 and 7 DF,  p-value: 1.912e-05
```

Question 3:Can you identify any optimal drug concentration ?

Answer 3: The optimal drug concentration can be identified with the help the ***graph*** and the ***histogram*** plotted and with the help of ***intercepts*** which we have taken while working out statistical test.

For MCF7

We can observe that the intercept value is 41.833

For Colo 205

We can observe that the intercept is 164.5915

For HeLa

We can observe that the intercept is -41.0605

For HMEC1

We can observe that the intercept is -56.0565

For HLF

We can observe that the intercept is 34.54791

Question 4: Can you suggest any changes in the experimental methodology in this study.

- To predict optimal concentration of the drug, Rather than comparing two concentrations (ie., 100ug/ml and 200 ug/ml). We can do more concentration according to the IC 50 of the particular drug.
- Since we are comparing the effect of the particular anti-cancer drug D1344320 on different types of cancer cells, all experimental setup should be uniform. In the case of given data, the number of cells treated in different cell lines is not equal.
- To validate the observed results experiment should be performed in triplicate.

Appendix

```
##For cell line MCF7

mcf7 <- as.data.frame(read.csv("MCF7.csv",sep=","))

head(mcf7)

length(mcf7$Time..hrs.)

length(mcf7$Number.of.Cells.0ug.mL.)

plot(mcf7$Time..hrs.,mcf7$Number.of.Cells.0ug.mL.,type="o",col="blue")

points(mcf7$Time..hrs.,mcf7$Number.of.Cells.100ug.mL.,type="o",col="orange")

points(mcf7$Time..hrs.,mcf7$Number.of.Cells.200ug.mL.,type="o",col="gray")

legend("topleft",legend=c("Number.of.Cells.0ug.mL.","Number.of.Cells.100ug.mL
.","Number.of.Cells.200ug.mL."),

      col=c("blue","orange","gray"),lty=c(1,1,1),cex = 0.5)



#linear regression

covariance <- cov(mcf7)

correlation <- cor(mcf7)

regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.

            + Number.of.Cells.100ug.mL.
```

```
            + Number.of.Cells.200ug.mL., data = mcf7)

covariance

correlation

regression

summary(regression)



##For Colo205 cell line

Colo205 <- as.data.frame(read.csv("Colo205.csv",sep=","))

head(Colo205)

length(Colo205$Time..hrs.)

length(Colo205$Number.of.Cells.0ug.mL.)

plot(Colo205$Time..hrs.,Colo205$Number.of.Cells.0ug.mL.,type="o",col="blue")

points(Colo205$Time..hrs.,Colo205$Number.of.Cells.100ug.mL.,type="o",col="orange")

points(Colo205$Time..hrs.,Colo205$Number.of.Cells.200ug.mL.,type="o",col="gray")

legend("bottomright",legend=c("Number.of.Cells.0ug.mL.","Number.of.Cells.100ug.mL.","Number.of.Cells.200ug.mL."),
```

```
      col=c("blue","orange","gray"),lty=c(1,1,1),cex = 0.5)


#linear regression

covariance <- cov(Colo205)

correlation <- cor(Colo205)

regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.

            + Number.of.Cells.100ug.mL.

            + Number.of.Cells.200ug.mL., data = Colo205)

covariance

correlation

regression

summary(regression)


##For HeLa cell line

HeLa <- as.data.frame(read.csv("HeLa.csv",sep=","))

head(HeLa)

length(HeLa$Time..hrs.)

length(HeLa$Number.of.Cells.0ug.mL.)
```

```
plot(HeLa$Time..hrs.,HeLa$Number.of.Cells.0ug.mL.,type="o",col="blue")

points(HeLa$Time..hrs.,HeLa$Number.of.Cells.100ug.mL.,type="o",col="orange"
)

points(HeLa$Time..hrs.,HeLa$Number.of.Cells.200ug.mL.,type="o",col="gray")

legend("bottomright",legend=c("Number.of.Cells.0ug.mL.","Number.of.Cells.100u
g.mL.","Number.of.Cells.200ug.mL."),

      col=c("blue","orange","gray"),lty=c(1,1,1),cex = 0.5)


#linear regression

covariance <- cov(HeLa)

correlation <- cor(HeLa)

regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.

            + Number.of.Cells.100ug.mL.

            + Number.of.Cells.200ug.mL., data = HeLa)

covariance

correlation

regression

summary(regression)
```

```
##For HMEC1 cell line

HMEC1 <- as.data.frame(read.csv("HMEC1.csv",sep=","))

head(HMEC1)

length(HMEC1$Time..hrs.)

length(HMEC1$Number.of.Cells.0ug.mL.)

plot(HMEC1$Time..hrs.,HMEC1$Number.of.Cells.0ug.mL.,type="o",col="blue")

points(HMEC1$Time..hrs.,HMEC1$Number.of.Cells.100ug.mL.,type="o",col="orange")

points(HMEC1$Time..hrs.,HMEC1$Number.of.Cells.200ug.mL.,type="o",col="gray")

legend("bottomright",legend=c("Number.of.Cells.0ug.mL.","Number.of.Cells.100ug.mL.","Number.of.Cells.200ug.mL."),

      col=c("blue","orange","gray"),lty=c(1,1,1),cex = 0.5)


#linear regression

covariance <- cov(HMEC1)

correlation <- cor(HMEC1)

regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.
```

```
        + Number.of.Cells.100ug.mL.
        + Number.of.Cells.200ug.mL., data = HMEC1)
covariance
correlation
regression
summary(regression)



##For HLF cell line
HLF <- as.data.frame(read.csv("HLF.csv",sep=","))
head(HLF)
length(HLF$Time..hrs.)
length(HLF$Number.of.Cells.0ug.mL.)
plot(HLF$Time..hrs.,HLF$Number.of.Cells.0ug.mL.,type="o",col="blue")
points(HLF$Time..hrs.,HLF$Number.of.Cells.100ug.mL.,type="o",col="orange")
points(HLF$Time..hrs.,HLF$Number.of.Cells.200ug.mL.,type="o",col="gray")
legend("bottomright",legend=c("Number.of.Cells.0ug.mL.","Number.of.Cells.100u
g.mL.","Number.of.Cells.200ug.mL."),
```

```
      col=c("blue","orange","gray"),lty=c(1,1,1),cex = 0.5)


#linear regression
covariance <- cov(HLF)
correlation <- cor(HLF)
regression <- lm(Time..hrs. ~ Number.of.Cells.0ug.mL.
                 + Number.of.Cells.100ug.mL.
                 + Number.of.Cells.200ug.mL., data = HLF)
covariance
correlation
regression
summary(regression)
```