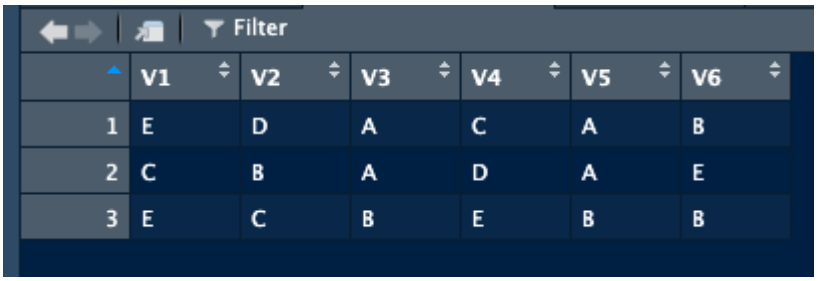
Problem Set - 1

[Layout and Analysis of Standard Designs]

1 . Construct the layout of a Completely Randomized Design with 5 treatments being replicated 4,5,3,2 and 4 times.

Solution :

```
rm(list=ls())
set.seed(123)
random_no = sample(seq(1,18,1),18,replace=F) ; random_no
data = c(rep("A",4),rep("B",5),rep("C",3),rep("D",2),rep("E",4))
treatment = data[random_no]
Design_Matrix = matrix(treatment,nrow=3,byrow=TRUE)
Design_Matrix
```



Filter

	V1	V2	V3	V4	V5	V6
1	E	D	A	C	A	B
2	C	B	A	D	A	E
3	E	C	B	E	B	B

2 . Construct the layout of a Randomized Block Design with 5 treatments and 4 blocks.

```
rm(list=ls())
set.seed(123)
Design_matrix = matrix(nrow=4, ncol=5)
treatment = c("A","B","C","D","E")
for(i in 1:4){
  Design_matrix[i,] = sample(treatment, 5, replace=FALSE)
}
print(block)

# Alternative Way to Do So

rm(list=ls())
set.seed(1234)
block_1 = sample(c("A","B","C","D","E"),5)
block_2 = sample(c("A","B","C","D","E"),5)
```

3. i) Construct a 4x4 Latin Square Design (LSD) with the treatments A,B,C,D.
ii) Construct a 5x5 Latin Square Design (LSD) with the treatments A,B,C,D,E.
iii) Construct a 6x6 Latin Square Design (LSD) with the treatments A,B,C,D,E,F.

Solution :

```
rm(list=ls())
create_LSD = function(n, treatments) {
  LSD = matrix(NA, nrow = n, ncol = n)
    for (i in 1:n) {
    for (j in 1:n) {
      LSD[i, j] = treatments[(i + j - 1) %% n + 1]
    }
  }
    return(LSD)
}
# Part- 1 [4x$ Order Matrix]
treatments = c("A", "B", "C", "D")
treatments = sample(treatments,length(treatments),replace=F)
M_4 = create_LSD(4, treatments)
M_4
# Part- 2 [5x5 Order Matrix]
treatments <- c("A", "B", "C", "D","E")
treatments = sample(treatments,length(treatments),replace=F)
M_5 = create_LSD(5, treatments)
M_5
# Part- 3 [5x5 Order Matrix]
treatments <- c("A", "B", "C", "D","E","F")
treatments = sample(treatments,length(treatments),replace=F)
M_6 = create_LSD(6, treatments)
M_6
```

4. A person wanting to purchase electric drills got quotations from five manufacturers. For selections, he wanted to conduct an experiment to estimate the time taken by each in making a hole in a metallic sheet. He marked 20 places on the sheet and applied five drills, labelled D_1, D_2, D_3, D_4, D_5 from each concern in randomly selected places to make the holes. The time for making each hole was recorded in seconds. The layout of the metallic sheet was as follows :

D1 (19)	D3 (22)	D4 (20)	D1 (20)
D5 (29)	D2 (24)	D5 (30)	D3 (24)
D2 (26)	D4 (25)	D1 (16)	D2 (22)
D5 (28)	D4 (25)	D5 (31)	D4 (28)
D4 (27)	D2 (16)	D2 (27)	D3 (20)

Analyze the model. Which drills create the difference in the time?

Solution :
This is an example of the One Way Anova model in CRD .
$y_{ij} = \mu + \tau_i + e_{ij}; i = 1(1)5, j = 1(1)n_i$

So , our Hypothesis of interest is then given by

$$H_0 : \tau_! = \tau_2 = \tau_3 = \tau_4 = \tau_5 \text{ against } H_1 : Not\ H_0$$

```
rm(list=ls())
time =
c(19,22,20,20,29,24,30,24,26,25,16,22,28,25,31,28,27,16,27,20)
drill =
c("D1","D3","D4","D1","D5","D2","D5","D3","D2","D4","D1","D2","D5"
,"D4","D5","D4","D4","D2","D2","D3")
data = data.frame(time,drill)
#Test for Treatment Effect
anova_table= summary(aov(time~as.factor(drill)));anova_table
summary(anova_table)
qf(0.95,4,15)
#test for pair significance
yi0.bar=aggregate(time~as.factor(drill),data=data,mean)$time;yi0.bar
n = aggregate(time~as.factor(drill),data=data,length)$time;n

#Or use this
yi0.bar <- by(data$time, data$drill, mean)
n <- by(data$time, data$drill, length)
```

```
t_obs = array(0);k=1;len=5
t_1 = array(0)
t_2 = array(0)
for(i in 1:(len-1))
  {
    for(j in (i+1):len)
      {
  t_obs[k] =
(yi0.bar[i]-yi0.bar[j])/sqrt(9.04*((1/n[i])+(1/n[j])))
  t_1[k] = i
  t_2[k] = j
  k=k+1
    }
  }
t_obs
qt(.975,15)
decision=ifelse(abs(t_obs)>qt(0.975,15),"Reject","Accept")
data.frame(t_1,t_2,T,decision)
```

5. To estimate the petrol consumption rates of different makes of cars for suitable average speed, five different cars each of four different makes were put on the road on 5 different days. The cars of a make ran with different speeds, 25, 35, 50, 60 and 70 mph on 5 different days. For each car, the number of miles per gallon of petrol was observed. The layout is given below.

Makes of Car	Speeds of Cars in Miles per hour (mph)				
	25	35	50	60	70
A	20.6	19.5	18.1	17.9	16.0
B	19.5	19.0	15.6	16.7	14.1
C	20.5	18.5	16.3	15.2	13.7
D	16.2	16.5	15.7	14.8	12.7

i) Identify the layout and the factors. Setup a suitable model for the data.

ii) Does the coverage differ significantly on the Makes of the Cars?
Test the same at 5% level of significance.

iii) If your answer to (ii) is in the negative, which makes account for the difference?

Solution :

```
rm(list=ls())
data = data.frame(
  Make = rep(c("A", "B", "C", "D"), each = 5),
  Speed = rep(c(25, 35, 50, 60, 70), times = 4),
  MPG = c(20.6, 19.5, 18.1, 17.9, 16.0,
```

```
          19.5, 19.0, 15.6, 16.7, 14.1,
          20.5, 18.5, 16.3, 15.2, 13.7,
          16.2, 16.5, 15.7, 14.8, 12.7)
)
anova_result = aov(MPG ~ as.factor(Speed) + as.factor(Make), data
= data)
summary(anova_result)

qf(0.95, 3, 12)
mean_MPG = aggregate(MPG ~ Make, data = data, mean)$MPG
mean_MPG

length = 4
k = 1
t_obs = array(0, dim = choose(len, 2))
for (i in 1:(len-1)) {
  for (j in (i+1):len) {
    t_obs[k] = (mean_MPG[i] - mean_MPG[j]) / sqrt(2 * 0.618 / 5)
    k = k + 1
  }
}
qt(.975, 12)
decision = ifelse(abs(t_obs) > qt(0.975, 12), "Reject", "Accept")
data.frame(t_obs, decision)
```

6. Read the "*OrchardSprays*" dataset in R. Identify the design and carry out the test for differential effects of the various constituents of orchard sprays in repelling honeybees. Interpret the result. If the hypothesis is rejected, find out which pairs of sprays cause the differential effects.

Solution :

```
rm(list=ls())
OrchardSprays
attach(OrchardSprays)
rowpos = as.factor(rowpos)
colpos = as.factor(colpos)
treatment = as.factor(treatment)
anova_table= aov(decrease~ rowpos+colpos+treatment);anova_table
summary(anova_table)
qf(.95,7,42)
```

```
yio.bar=aggregate(decrease~as.factor(treatment),data=OrchardSprays
,mean)$decrease;yio.bar
len=8
k=1
t_obs=array(0)
for(i in 1:(len-1))
  {
  for(j in (i+1):len)
    {
    t_obs[k]=(yio.bar[i]-yio.bar[j])/sqrt(2*381/8)
  k=k+1
    }
  }
t_obs
qt(.975,42)
decision = ifelse(abs(t_obs)>qt(0.975,42),"Reject","Accept")
data.frame(t_obs,decision)
detach(OrchardSprays)
```

7. i) Consider the following ANOVA table of an RBD.

Sources of Variation	**d.f.**	**SS**
Blocks	3	3.2947
Treatments	4	68.8917
Error	12	3.2767

Estimate the efficiency of the RBD with respect to a CRD and comment.

Solution :

Converting this CRD Setup under *Uniformity Trials*, we get -

S.V	d.f.	SS	MS	
Blocks	3	3.2947	1.098233	
Error	16	72.1587	4.5099	MSE_{CRD}
Total	19	75.4631		

Converting this CRD Setup into a Comparable LSD and under *Uniformity Trials*, we get -

S.V	d.f.	SS	MS	
Error	19	75.4631	3.97164	MSE_{RBD}
Total	19	75.4631		

Now, the efficiency of the RBD Setup wrt. CRD setup is provided by :

$$\frac{Eff_{CRD}}{Eff_{RBD}} = \frac{MSE_{RBD}}{MDE_{CRD}} = \frac{3.97164}{4.5099} = 0.8806 < 1 \Rightarrow Eff_{CRD} < Eff_{RBD}$$

So, RBD is more efficient than the CRD.

(ii) Consider the following ANOVA table of a LSD.

Sources of Variation	**d.f.**	**SS**
Rows	7	102.19
Columns	7	84.24
Treatments	7	513.79
Error	42	92.00

Estimate the efficiency of the LSD (once with rows as blocks and once with columns as blocks) with respect to an RBD and comment.

Solution **:**
Converting this LSD under uniformity trails, we get that :

S.V	d.f.	SS	
Rows	7	102.19	
Treatments	7	513.79	
Error	42	176.24	$\frac{176.24}{42} = MSE_{LSD}$
Total	63	792.03	

Converting this LSD into an RBD Setup and under uniformity trails, we get that :

S.V	d.f.	SS	
Rows	7	102.19	
Error	56	689.84	$\frac{689.84}{56} = MSE_{RBD}$
Total	63	792.03	

$$\frac{Eff_{RBD}}{Eff_{LSD}} = \frac{MSE_{LSD}}{MDE_{RBD}} = \frac{4.1961904}{12.31857} = 0.340639 < 1 \Rightarrow Eff_{RBD} < Eff_{LSD}$$

So, LSD is more efficient than the RBD.

Problem Set - 2

[Missing Plot Technique]

1. The following data were obtained from an experiment using the treatment Blitox, Dithane, Brestan and Control. After sowing rhizomes of the mat-grass in four plots in each of the three villages, the above four treatments were applied at random to the plots in a village after 30 days of sowing. Unfortunately the yield of the third treatment in the third village was lost. The available yields in gm of 1 rg ft cutting per plot after 120 days are given below.

   i) Estimate the missing observation based on the available data.
   ii) Analyse the data to find out if there are any significant treatment effects.
   iii) Test if the treatments Blitox and Brestan differ significantly on the yield. Also find an estimate of the variance of the estimate of the treatment difference.
   iv) Obtain an estimate of the average variance of the elementary contrasts.

Treatment	village		
	1	2	3
Blitox	678	510	531
Dithane	703	689	611
Brestan	736	574	*
Control	556	510	500

Solution :

```
# Part - 1
rm(list=ls())
names = c("Blitox","Dithane","Breastan","Control")
data = c(678,703,736,556,510,689,574,510,531,611,594.66,500)
treatment = rep(c(1,2,3,4),3)
block = c(1,1,1,1,2,2,2,2,3,3,3,3)
model = aov(data~as.factor(treatment)+as.factor(block))
summary(model)
SSE_x_hat_hat = 10286 ; SSE_x_hat_hat
x_hat_hat = 1642/3 ; x_hat_hat
data_HO = c(678,703,736,556,510,689,574,510,531,611,x_hat_hat,500)
model_HO = aov(data_HO~as.factor(treatment)+as.factor(block))
summary(model_HO)
qf(0.95,3,5)

# Part - 2
SSE = 10286
```

```
SSE_HO = 35022+11406
v = 4 ; b=3
F_obs = ((SSE_HO - SSE)/(v-1))/(SSE/((v-1)*(b-1)-1)) ; F_obs
qf(0.95,3,5)
if(F_obs>qf(0.95,3,5)){
  print("Reject")
  }else{
    print("Accept")
  }
sum(data)

# Part - 3
y01_bar = (678+510+531)/3 ; y01_bar
y03_bar = (736+574+594.66)/3 ; y03_bar
MSE = 10286/5
E_hat = (y01_bar - y03_bar) ; E_hat
V_hat = MSE*((2/b)+(v/(b*(b-1)*(v-1)))) ; V_hat
t = E_hat/sqrt(V_hat) ; t


# Part - 4
Avg_variance = (6*(MSE/b) + 3*MSE*(2/b + v/(b*(v-1)*(b-1))))/6
Avg_variance
```

2. The following data were obtained from a cow-feeding experiment with 5 feeds using a Latin Square Design.

i) Estimate the missing observation based on the available data.

ii) Analyse the data to find out if there are any significant treatment effects.

iii) If the treatments are found to be significant, obtain which pairs cause the differential effects.

iv) Obtain an estimate of the average variance of the elementary contrasts.

A (30)	B (35)	C (47)	D(50)	E (38)
B (36)	C (42)	D (48)	E (*)	A (32)
C (41)	D (52)	E (40)	A (34)	B (38)
D (54)	E (40)	A (35)	B (34)	C (45)
E (40)	A (34)	B (38)	C (42)	D (50)

Solution :

Problem Set - 3

H/Design/Sem6/Problem Set 3 **Date:**

Factorial Experiments

1. A bacteriologist is interested in the effect of two different culture mediums and two different times on the growth of a particular virus. She performs six replicates of a 2^2 design, making the runs in random order. Analyse the data that follow and draw appropriate conclusions.

Time	Culture Medium			
	1		2	
12 hr	21	22	25	26
	23	28	24	25
	20	26	29	27
18 hr	37	39	31	34
	38	38	29	33
	35	36	30	35

Solution :

```
rm(list=ls())
culture_medium = rep(c(1, 2), each=12)
time = rep(c(12, 18), times=12)
virus_growth = c(21, 22, 25, 26, 23, 28, 24, 25, 20, 26, 29, 27,
                 37, 39, 31, 34, 38, 38, 29, 33, 35, 36, 30, 35)

data = data.frame(culture_medium, time, virus_growth)
data$culture_medium = as.factor(data$culture_medium)
data$time = as.factor(data$time)
anova_results = aov(virus_growth ~ culture_medium * time,
data=data)
summary(anova_results)
```

2 . An experiment was performed to improve the yield of a chemical process. Four factors were selected and three replicates of a randomized block experiment were run. The results are shown in the following table. Analyse the data and draw appropriate conclusions:

	Replicate				Replicate		
	1	2	3		1	2	3
1	90	93	98	d	98	95	100
a	74	78	85	ad	72	76	82
b	81	85	88	bd	87	83	91
ab	83	80	84	abd	85	86	86
c	77	78	82	cd	99	90	98
ac	81	80	85	acd	79	75	81
bc	88	82	88	bcd	87	84	86
abc	73	70	79	abcd	80	80	85

Solution :

```
rm(list=ls())
data =
c(90,93,98,74,78,85,81,85,88,83,80,84,77,78,82,81,80,85,88,82,88,7
3,70,79,98,95,100,72,76,82,87,83,91,85,86,86,99,90,98,79,75,81,87,
84,86,80,80,85)
Block = factor(rep(1:3,16))
A = factor(rep(rep(c("1","a"),each=3),8))
B = factor(rep(rep(c("1","b"),each=6),4))
C = factor(rep(rep(c("1","c"),each=12),2))
D = factor(rep(rep(c("1","d"),each=24),1))

Chemical.process = data.frame(data,Block,A,B,C,D)
View(Chemical.process)
Chem.Pro.aov = aov(data ~ Block+A*B*C*D,data=Chemical.process)
summary(Chem.Pro.aov)
names(Chem.Pro.aov)
qf(0.95,1,30)
```

H/Design/Sem6/Problem Set 6 **Date:**

Confounding in Factorial Experiments

1. An engineer is interested in the effect of cutting speed (A), tool geometry (B), and cutting angle (C) on the life of a machine tool. Two levels of each factor are chosen and three replicates of a 2^3 factorial design are run. The results follow:

	Replicate		
	1	2	3
(1)	22	31	25
a	32	43	29
B	35	34	50
ab	55	47	46
C	44	45	38
ac	40	37	36
bc	60	50	54
abc	39	41	47

Suppose ABC is confounded in Replicate 1, AB in replicate 2 and BC in replicate 3. Construct and analyse the design.

Solution:

```
rm(list=ls())
data=c(22,32,35,55,44,40,60,39,31,43,34,47,45,37,50,41,25,29,50,46
,38,36,54,47)
blocks=as.factor(c(1,2,2,1,2,1,1,2,3,4,4,3,3,4,4,3,5,5,6,6,6,6,5,5
))
A=rep(0:1,times=12)
B=rep(0:1,each=2,times=6)
C=rep(0:1,each=4,times=3)
anova.fit=aov(data~A*B*C+blocks)
summary(anova.fit)

data.frame(data,blocks)
```

2. An experiment was conducted to investigate the effects of 5 factors on the yield of penicillin in surface culture experiments. The factors and their levels were corn steep liquor (Factor A at 2% and 3% strength), lactose (Factor B at 2% and 3% strength), precursor (Factor C at 0% and 0.05% strength), sodium nitrate (Factor D at 0% and 0.3% strength) and glucose (Factor E at 0% and 0.5% strength). A complete replicate of the experiment yields the following data Construct a design in four blocks with ABCD, CDE and consequently ABE confounded with the blocks. Test for the main effects and two factor interaction effects assuming that the unconfounded three, four and five factor interactions estimate the error.

00000 (142)	00001 (106)	00010 (148)	00011 (101)	00100 (185)	00101 (113)	00110 (200)	00111 (130)
01000 (129)	01001 (88)	01010 (146)	01011 (140)	01100 (200)	01101 (166)	01110 (215)	01111 (145)
10000 (114)	10001 (106)	10010 (108)	10011 (114)	10100 (162)	10101 (88)	10110 (164)	10111 (83)
11000 (109)	11001 (98)	11010 (195)	11011 (72)	11100 (79)	11101 (172)	11110 (118)	11111 (110)

Solution :

```
rm(list=ls())
data=c(142,106,148,101,185,113,200,130,129,88,146,140,200,166,215,
145,114,106,108,114,162,88,164,83,109,98,195,72,79,172,118,110)
sum(data)
blocks=as.factor(c(1,4,3,2,3,2,1,4,2,3,4,1,4,1,2,3,2,3,4,1,4,1,2,3
,1,4,3,2,3,2,1,4))
sum(ifelse(blocks==3,1,0))
A=rep(0:1,each=16,times=1)
B=rep(0:1,each=8,times=2)
C=rep(0:1,each=4,times=4)
D=rep(0:1,each=2,times=8)
E=rep(0:1,each=1,times=16)
anova.fit=aov(data~blocks+A*B*C*D*E)
summary(anova.fit)

anova.fit2=aov(data~blocks+A*B+A*C+A*D+A*E+B*C+B*D+B*E+C*D+C*E+D*E
)
summary(anova.fit2)
#Alternative
anova.fit2=aov(data~blocks+(A+B+C+D+E)^2)
summary(anova.fit2)
#Alternative for 4,5-factor not needed
anova.fit2=aov(data~blocks+(A+B+C+D+E)^3)
summary(anova.fit2)
```

3. Consider the 2^6 design in eight blocks of eight runs each, with ABCD, AEF and ACDE as the independent effects chosen to be confounded with blocks. Generate the design. Find the other effects confounded with the blocks.