1. (a) The model in this problem is $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ with $\boldsymbol{\beta} = (\mu, \alpha_1, \alpha_2, \beta_1, \beta_2)'$ and

$$\boldsymbol{X} = \left(\mathbf{1}_{20\times 1}, \begin{bmatrix} \mathbf{1}_{10\times 1} \\ \mathbf{0}_{10\times 1} \end{bmatrix}, \begin{bmatrix} \mathbf{0}_{10\times 1} \\ \mathbf{1}_{10\times 1} \end{bmatrix}, \begin{bmatrix} \mathbf{1}_{2\times 1} \\ \mathbf{0}_{8\times 1} \\ \mathbf{1}_{6\times 1} \\ \mathbf{0}_{4\times 1} \end{bmatrix}, \begin{bmatrix} \mathbf{0}_{2\times 1} \\ \mathbf{1}_{8\times 1} \\ \mathbf{0}_{6\times 1} \\ \mathbf{1}_{4\times 1} \end{bmatrix} \right)$$

Because $rank(\boldsymbol{X}) = 3$, one possible full rank matrix is

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = \left(\mathbf{1}_{20\times 1}, \begin{bmatrix} \mathbf{1}_{10\times 1} \\ \mathbf{0}_{10\times 1} \end{bmatrix}, \begin{bmatrix} \mathbf{1}_{2\times 1} \\ \mathbf{0}_{8\times 1} \\ \mathbf{1}_{6\times 1} \\ \mathbf{0}_{4\times 1} \end{bmatrix} \right)$$

We can compute orthogonal columns using Gram-Schmidt Orthogonalization method:

$$\begin{aligned} \boldsymbol{w}_1 &= \boldsymbol{x}_1; \\ \boldsymbol{w}_2 &= (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{w}_1})\boldsymbol{x}_2 = \frac{1}{2} \begin{bmatrix} \mathbf{1}_{10\times 1} \\ -\mathbf{1}_{10\times 1} \end{bmatrix} \\ \boldsymbol{w}_3 &= (\boldsymbol{I} - \boldsymbol{P}_{[\boldsymbol{w}_1, \boldsymbol{w}_2]})\boldsymbol{x}_3 = \begin{bmatrix} 0.8 \cdot \mathbf{1}_{2\times 1} \\ -0.2 \cdot \mathbf{1}_{8\times 1} \\ 0.4 \cdot \mathbf{1}_{6\times 1} \\ -0.6 \cdot \mathbf{1}_{4\times 1} \end{bmatrix} \end{aligned}$$

So a model matrix $\boldsymbol{W}$ with orthogonal columns is $\boldsymbol{W} = (\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$.

(b) Compute $\boldsymbol{P}_{\boldsymbol{W}}\boldsymbol{y}$:

$$\begin{aligned} (\boldsymbol{W}'\boldsymbol{W})^{-1} &= \left[\begin{pmatrix} \boldsymbol{w}_1' \\ \boldsymbol{w}_2' \\ \boldsymbol{w}_3' \end{pmatrix} (\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3) \right]^{-1} = \begin{bmatrix} \boldsymbol{w}_1'\boldsymbol{w}_1 & & \\ & \boldsymbol{w}_2'\boldsymbol{w}_2 & \\ & & \boldsymbol{w}_3'\boldsymbol{w}_3 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 20 & & \\ & 5 & \\ & & 4 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{20} & & \\ & \frac{1}{5} & \\ & & \frac{1}{4} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \boldsymbol{W}'\boldsymbol{y} &= \begin{pmatrix} \boldsymbol{w}_1' \\ \boldsymbol{w}_2' \\ \boldsymbol{w}_3' \end{pmatrix} \cdot \boldsymbol{y} = \begin{bmatrix} \boldsymbol{w}_1'\boldsymbol{y} \\ \boldsymbol{w}_2'\boldsymbol{y} \\ \boldsymbol{w}_3'\boldsymbol{y} \end{bmatrix} \\ &= \begin{bmatrix} \sum_i \sum_j \sum_k y_{ijk} \\ \frac{1}{2}(\sum_j \sum_k y_{1jk} - \sum_j \sum_k y_{2jk}) \\ 0.8 \cdot \sum_k y_{11k} - 0.2 \cdot \sum_k y_{12k} + 0.4 \cdot \sum_k y_{21k} - 0.6 \cdot \sum_k y_{22k} \end{bmatrix} \\ &= \begin{bmatrix} 100 \\ -4 \\ 6.4 \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
\boldsymbol{P_W y} &= \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{y} \\
&= (\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3) \begin{bmatrix} \frac{1}{20} & & \\ & \frac{1}{5} & \\ & & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 100 \\ -4 \\ 6.4 \end{bmatrix} \\
&= 5\boldsymbol{w}_1 - 0.8\boldsymbol{w}_2 + 1.6\boldsymbol{w}_3 \\
&= \begin{bmatrix} 5.88 \cdot \mathbf{1}_{2\times 1} \\ 4.28 \cdot \mathbf{1}_{8\times 1} \\ 6.04 \cdot \mathbf{1}_{6\times 1} \\ 4.44 \cdot \mathbf{1}_{4\times 1} \end{bmatrix}
\end{aligned}$$

(c) The Type II sum of squares for factor B is $S(B|1, A) = \boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}$ where $\boldsymbol{P}_3\boldsymbol{y} = \boldsymbol{P_W y}$ and from part (b)

$$\begin{aligned}
\boldsymbol{P}_2\boldsymbol{y} &= \boldsymbol{P}_{[\boldsymbol{w}_1, \boldsymbol{w}_2]}\boldsymbol{y} \\
&= (\boldsymbol{w}_1, \boldsymbol{w}_2) \begin{bmatrix} \frac{1}{20} & \\ & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 100 \\ -4 \end{bmatrix} \\
&= 5\boldsymbol{w}_1 - 0.8\boldsymbol{w}_2 \\
&= \begin{bmatrix} 4.6 \cdot \mathbf{1}_{10\times 1} \\ 5.4 \cdot \mathbf{1}_{10\times 1} \end{bmatrix}
\end{aligned}$$

so

$$\begin{aligned}
S(B|1, A) &= \boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{P}_3\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{P}_2\boldsymbol{y} = ||\boldsymbol{P}_3\boldsymbol{y}||^2 - ||\boldsymbol{P}_2\boldsymbol{y}||^2 \\
&= (5.88^2 \times 2 + 4.28^2 \times 8 + 6.04^2 \times 6 + 4.44^2 \times 4) - (4.6^2 \times 10 + 5.4^2 \times 10) \\
&= 10.24
\end{aligned}$$

2. (a) For the jth woman treated with the ith drug,

$$\begin{aligned}
\boldsymbol{W} &= \mathrm{Var}(y_{ij1}, y_{ij2}, y_{ij3}, y_{ij4})' \\
&= \begin{bmatrix} \sigma_w^2 + \sigma_e^2 & \sigma_w^2 & \sigma_w^2 & \sigma_w^2 \\ \sigma_w^2 & \sigma_w^2 + \sigma_e^2 & \sigma_w^2 & \sigma_w^2 \\ \sigma_w^2 & \sigma_w^2 & \sigma_w^2 + \sigma_e^2 & \sigma_w^2 \\ \sigma_w^2 & \sigma_w^2 & \sigma_w^2 & \sigma_w^2 + \sigma_e^2 \end{bmatrix} \\
&= \sigma_w^2 \mathbf{1}\mathbf{1}'_{4\times 4} + \sigma_e^2 \boldsymbol{I}_{4\times 4}.
\end{aligned}$$

We know that $\mathrm{Var}(\boldsymbol{y})$ is block diagonal with blocks $\boldsymbol{W}$. There are a total of $3 \cdot 5 = 15$ blocks, so that

$$\mathrm{Var}(\boldsymbol{y}) = \boldsymbol{I}_{15\times 15} \otimes \boldsymbol{W} = \boldsymbol{I}_{15\times 15} \otimes (\sigma_w^2 \mathbf{1}\mathbf{1}'_{4\times 4} + \sigma_e^2 \boldsymbol{I}_{4\times 4}).$$

(b) The null hypothesis of no drug-by-time interactions is $H_0 : \mu_{ij} - \mu_{ij*} = \mu_{i*j} - \mu_{i*j*}$ for all $i \neq i*$ and $j \neq j*$. The test statistic $F = 7.12$ on $(6, 36)$ degrees of freedom with $p < 0.001$. We reject the null hypothesis and conclude that there is significant evidence for drug-by-time interactions on heart rate.

(c) The null hypothesis for testing the same mean heart rate 15 minutes after treatment for all three drugs is $H_0 : \mu_{14} = \mu_{24} = \mu_{34}$. The test statistic $F = 1.11$ on $(2, 17.1)$ degrees of freedom with $p = 0.352 > 0.05$. We fail to reject the null hypothesis and conclude that there is no significant evidence for the same mean heart rate 15 minutes after treatment for all three drugs.

(d) An approximate 95% confidence interval for $\mu_{14} - \mu_{24}$ is $(-13.77, 2.57)$ with df = 17.1 by the SAS code below.

Note: for part (c-d), df = 17.1 was computed by Cochran-Satterthwaite since it is for the difference between simple effects with different whole-plot factors. The easiest way to get this is to use SAS with ddfm=satterthwaite option.

SAS code:

```
proc import datafile="./HeartRate.txt"
 dbms=TAB replace out=d;
run;
proc print data=d (obs=14);
run;
proc mixed;
class woman drug time;
model y=drug time drug*time /ddfm = satterthwaite;
random woman(drug);
contrast "same mean for drug A B C at 15 min"
drug 1 -1 0 drug*time 0 0 0 1 0 0 0 -1 0 0 0 0 ,
drug 1 0 -1 drug*time 0 0 0 1 0 0 0 0 0 0 0 -1 ;
estimate "drug A - drug B at 15 min"
 drug 1 -1 0 drug*time 0 0 0 1 0 0 0 -1 0 0 0 0 /cl;
run;
```

R code:

```
library(MASS)
library(nlme)
library(dplyr)

> d <- read.table("http://dnett.github.io/S510/HeartRate.txt", header = T)
> d$t <- factor((d$time + 5) / 5)   # Levels of time.
> fit <- lme(y ~ drug * t, random = ~ 1 | woman, data = d)
> anova(fit)
            numDF denDF   F-value p-value
(Intercept)     1    36 2900.7782  <.0001
drug            2    12    1.3517  0.2955
t               3    36   10.2159  0.0001
drug:t          6    36    7.1153  <.0001
>
> # Function from Dr. Nettletons Notes
> test=function(lmout,C,d=0,df){
+    b=fixed.effects(lmout)
+    V=vcov(lmout)
+    dfn=nrow(C)
+    Cb.d=C %*% b - d
+    Fstat=drop(t(Cb.d)%*%solve(C%*%V%*%t(C))%*%Cb.d/dfn)
+    pvalue=1-pf(Fstat,dfn,df)
```

```
+    cbind(Fstat=Fstat,pvalue=pvalue)
+ }
> C1 <- matrix(c(0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,
+                0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1), byrow=T,nrow = 2)
> test(fit,C1,df=17.1) # df computed by Cochran-Satterthwaite via SAS.
        Fstat    pvalue
[1,] 1.109573 0.3523058
>
> # Function from Dr. Nettletons Notes
> ci <- function(lmeout, C, df, a = 0.05) {
+   b = fixed.effects(lmeout)
+   V = vcov(lmeout)
+   Cb = C %*% b
+   se = sqrt(diag(C %*% V %*% t(C)))
+   tval = qt(1 - a / 2, df)
+   low = Cb - tval * se
+   up = Cb + tval * se
+   m = cbind(C, Cb, se, low, up)
+   dimnames(m)[[2]] = c(paste("c", 1 : ncol(C), sep = ""),
+                        "estimate", "se", paste(100 * (1 - a), "% Conf.", sep = ""), "limits")
+   return(m)
+ }
> C2 <- matrix(c(0, -1, 0, 0, 0, 0, 0, 0, 0, 0, -1, 0), nrow = 1)
> ci(fit, C2, 17.1) # df computed by Cochran-Satterthwaite via SAS.
     c1 c2 c3 c4 c5 c6 c7 c8 c9 c10 c11 c12 estimate       se
[1,]  0 -1  0  0  0  0  0  0  0   0  -1   0     -5.6 3.872564
     95% Conf.   limits
[1,] -13.76676 2.566758
```

3. (a) Under a compound symmetry assumption,

$$\boldsymbol{W} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix},$$

where the REML estimates for the heart rate data are $\hat{\sigma} = 6.12$ and $\hat{\rho} = 0.777$. $(\hat{\sigma}_s^2 = 29.13, \hat{\sigma}_e^2 = 8.36)$

(b) Using R, AIC =-2(-144.9602)+2(14)=317.92 and BIC =-2(-144.9602)+(14)log(48)= 344.12. Using SAS, AIC =-2(-144.9602)+2(2)= 293.9 and BIC = -2(-144.9602)+(2)log(15)=295.3.

(c) Under an AR(1) assumption,

$$\boldsymbol{W} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix},$$

where the REML estimates for the heart rate data are $\hat{\sigma} = 6.00$ and $\hat{\rho} = 0.828$.

(d) Using R, AIC =-2(-142.9713)+2(14)=313.94 and BIC=-2(-142.9713)+(14)log(48)= 340.14. Using SAS, AIC =-2(-142.9713)+2(2)= 289.9 and BIC =-2(-142.9713)+(2)log(15)= 291.4.

(e) Under a general symmetry assumption,

$$\boldsymbol{W} = \sigma^2 \begin{bmatrix} 1 & \rho_{12}\delta_2 & \rho_{13}\delta_3 & \rho_{14}\delta_4 \\ \rho_{12}\delta_2 & \delta_2^2 & \rho_{23}\delta_2\delta_3 & \rho_{24}\delta_2\delta_4 \\ \rho_{13}\delta_3 & \rho_{23}\delta_2\delta_3 & \delta_3^2 & \rho_{34}\delta_3\delta_4 \\ \rho_{14}\delta_4 & \rho_{24}\delta_2\delta_4 & \rho_{34}\delta_3\delta_4 & \delta_4^2 \end{bmatrix},$$

where the REML estimates for the heart rate data are

$$\begin{aligned} &\hat{\sigma} = 6.10, \\ &\hat{\delta}_2 = 1.08, \quad \hat{\delta}_3 = 0.995, \quad \hat{\delta}_4 = 0.928, \\ &\hat{\rho}_{12} = 0.850, \quad \hat{\rho}_{13} = 0.889, \quad \hat{\rho}_{14} = 0.625, \\ &\hat{\rho}_{23} = 0.870, \quad \hat{\rho}_{24} = 0.631, \quad \hat{\rho}_{34} = 0.794. \end{aligned}$$

(f) Using R, AIC =-2(-139.424)+2(22)=322.85 and BIC =-2(-139.424)+(22)log(48)= 364.01. Using SAS, AIC =-2(-139.424)+2(10)= 298.8 and BIC =-2(-139.424)+(10)log(15)= 305.9.

(g) The model with an AR(1) correlation structure has the smallest AIC and BIC of the three (regardless of whether you used R or SAS). Consequently, the AR(1) correlation structure is preferred for this dataset.

(h) There are several ways to find a 95% confidence interval for $\mu_{14} - \mu_{24}$ using the model with an AR(1) correlation structure. In question 2 (d), we used a split-plot design to get a confidence interval, for which it is clear that we should compute the degrees of freedom using Cochran-Satterthwaite. However, it is less clear for the model using AR(1). Regardless of which degrees of freedom you use, you should have $\widehat{\mu_{14} - \mu_{24}} = -5.6$ with $\sqrt{\widehat{\text{Var}}(\widehat{\mu_{14} - \mu_{24}})} = 3.795$.

We can use the Cochran-Satterthwaite method in SAS by specifying the "ddfm = satterthwaite" option, which gives the interval $(-13.54, 2.34)$ based on df = 19.2. The default "ddfm" method in SAS uses df = 36, which gives the interval $(-13.30, 2.10)$.

In R, gls computes df $= n - \text{rank}(X) = 48$, which leads to the interval $(-13.23, 2.03)$.

Of these intervals, I would prefer the one where the degrees of freedom are computed by Cochran-Satterthwaite because it is the widest and hence the most conservative in terms of inference about the value of $\mu_{14} - \mu_{24}$.

SAS code:

```
proc mixed;
     class woman drug time;
     model y = drug time drug*time;
```

```
     repeated time / subject = woman type = cs r rcorr;
run;
proc mixed;
     class woman drug time;
     model y = drug time drug*time / ddfm = satterthwaite;
     repeated time / subject = woman type = ar(1) r rcorr;
     estimate 'drug A - drug B at 15 minutes'
          drug 1 -1 0 drug * time 0 0 0 1   0 0 0 -1   0 0 0 0 / cl;
run;
proc mixed;
     class woman drug time;
     model y = drug time drug*time;
     repeated time / subject = woman type = un  r rcorr;
run;
```

R code:

```
attach(d)
woman <- as.factor(woman)
drug <- as.factor(drug)
time <- as.factor(time)
model.cs <- gls(y ~ drug * time,
                correlation = corCompSymm(form = ~1 | woman),
                method = "REML")
model.ar <- gls(y ~ drug * time,
                correlation = corAR1(form = ~1 | woman),
                method = "REML")
model.sy <- gls(y ~ drug * time,
                correlation = corSymm(form = ~1 | woman),
                weight = varIdent(form = ~1 | time),
                method = "REML")
summary(model.cs)
getVarCov(model.cs)
summary(model.ar)
getVarCov(model.ar)
summary(model.sy)
getVarCov(model.sy)
ci.gls <- function(lmeout, C, df, a = 0.05) {
  b = coef(lmeout)
  V = vcov(lmeout)
  Cb = C %*% b
  se = sqrt(diag(C %*% V %*% t(C)))
  tval = qt(1 - a / 2, df)
  low = Cb - tval * se
  up = Cb + tval * se
  m = cbind( Cb, se, low, up)
  dimnames(m)[[2]] = c("estimate", "se", paste(100 * (1 - a), "% Conf.", sep = ""), "limits")
  return(m)
}
ci.gls(model.ar, C2, 19.2) # Cheated and took Cochran-Satterthwaite df value from SAS.
ci.gls(model.ar, C2, 36) # Default df method in SAS.
ci.gls(model.ar, C2, 48) # Default df method in R.
```

4. The model we used in this problem is $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zs} + \boldsymbol{e}$ where $\boldsymbol{y} = (\boldsymbol{y}_1', \boldsymbol{y}_2', ..., \boldsymbol{y}_{50}')', \boldsymbol{e} = (\boldsymbol{e}_1', \boldsymbol{e}_2', ..., \boldsymbol{e}_{50}')', \boldsymbol{s} = (s_1, s_2, ..., s_{50})'$, $\boldsymbol{y}_1 = (y_{11}, y_{12})'$, $\boldsymbol{e}_1 = (e_{11}, e_{12})'$, $\boldsymbol{y}_i = (y_{i1}, y_{i2}, y_{i3})'$ and $\boldsymbol{e}_i = (e_{i1}, e_{i2}, e_{i3})'$ for $i = 2, ..., 50$.

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ & \vdots & \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ \boldsymbol{Z} = diag(\boldsymbol{1}_2, \boldsymbol{I}_{49} \otimes \boldsymbol{1}_3), \ \boldsymbol{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} \text{ and } \boldsymbol{\epsilon} = \boldsymbol{Zs} + \boldsymbol{e} \sim N\left(\boldsymbol{0}_{149\times 1}, \boldsymbol{\Sigma}\right).$$

$Var(\boldsymbol{y}) = \boldsymbol{\Sigma} = diag(\boldsymbol{W}_1, \boldsymbol{I}_{49} \otimes \boldsymbol{W})$ is a block diagonal matrix with, for $i = 2, ..., 50$,

$$Var\left(\boldsymbol{y}_1\right) = \boldsymbol{W_1} = \sigma_s^2 \boldsymbol{1}_2 \boldsymbol{1}_2' + \sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & \delta_2^2 \end{bmatrix},$$

$$VAR\left(\boldsymbol{y}_i\right) = \boldsymbol{W} = \sigma_s^2 \boldsymbol{1}_3 \boldsymbol{1}_3' + \sigma^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & \delta_2^2 & 0 \\ 0 & 0 & \delta_3^2 \end{bmatrix}.$$

(a) Under a model above, the REML estimates for the exam score data are

$$\hat{\sigma}_s = 13.43525, \quad \hat{\sigma} = 7.933829 \quad \hat{\delta}_2 = 0.978757, \quad \hat{\delta}_3 = 0.522279$$
$$\hat{\sigma}_s^2 = 180.5059, \quad \hat{\sigma}_1^2 = \hat{\sigma}^2(1) = 62.9456, \quad \hat{\sigma}_2^2 = \hat{\sigma}^2\hat{\delta}_2^2 = 60.2914, \quad \hat{\sigma}_3^2 = \hat{\sigma}^2\hat{\delta}_3^2 = 17.1676.$$

```
> library(nlme)
> d=read.table("http://dnett.github.io/S510/ExamScores.txt",
+                header = T, colClasses = c("factor","factor","numeric"))
> output=lme(score ~ 0 + exam, random = ~ 1 | student,
+         weights = varIdent(form = ~ 1 | exam), data = d)
> output
Linear mixed-effects model fit by REML
  Data: d
  Log-restricted-likelihood: -552.6604
  Fixed: score ~ 0 + exam
   exam1    exam2    exam3
46.84000 58.16000 67.25471

Random effects:
 Formula: ~1 | student
        (Intercept) Residual
StdDev:    13.43525 7.933829
```

```
Variance function:
 Structure: Different standard deviations per stratum
 Formula: ~1 | exam
 Parameter estimates:
       1        2        3
1.000000 0.978757 0.522279
Number of Observations: 149
Number of Groups: 50
```

(b) The eBLUP for student 1's exam 3 score is 83.7345 by the R code below.

```
> fixef(output)[3]+ranef(output)[1,]
   exam3
83.73545
```