

# MS2

Likhitha Mandapati

2023-04-30

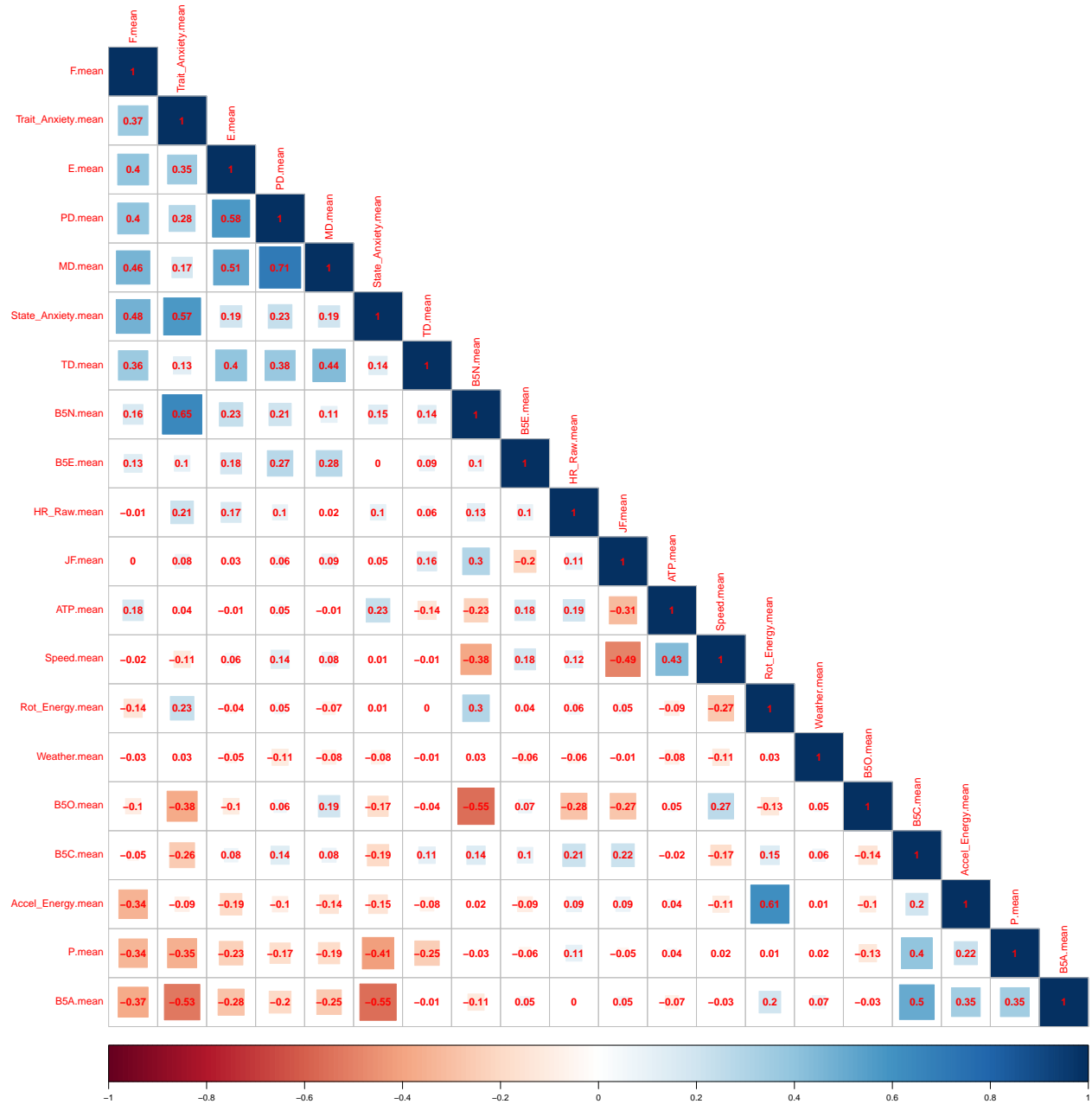
## Contents

<b>Exploratory analysis</b>	<b>2</b>
<b>Data corealtion plot</b>	<b>2</b>
<b>HR_Raw Full model</b>	<b>3</b>
Full model summary . . . . .	3
Full Model AIC . . . . .	4
Full model Plots . . . . .	5
<b>Backward Elimination (BE)</b>	<b>8</b>
BE steps . . . . .	8
BE summary . . . . .	9
BE AIC . . . . .	10
<b>Forward Elimination (FE)</b>	<b>11</b>
FE steps . . . . .	11
FE summary . . . . .	11
FE AIC . . . . .	11
<b>Both Direction Elimination (BDE)</b>	<b>12</b>
BDE steps . . . . .	12
BE summary . . . . .	12
BE AIC . . . . .	12
<b>Optimal model</b>	<b>13</b>
Optimal model summary . . . . .	13
Optimal model plots . . . . .	14
Predictor plot . . . . .	14
<b>Conclusion</b>	<b>16</b>

## Exploratory analysis

- In MS1, we performed exploratory data analysis (EDA) for all the predictors in the dataset. Based on the insights we gained from the EDA, we can now build a more effective and informed linear mixed model in MS2.

## Data corealtion plot



- MD - PD and ATP - RTP are highly correlated, so we should consider only one among each of them.
- We consider MD and ATP.

## HR\_Raw Full model

```
full <- lmer(HR_Raw.mean ~ 1 + Weather.mean + JF.mean +Speed.mean +ATP.mean +
  Accel_Energy.mean + Rot_Energy.mean + State_Anxiety.mean +
  B5A.mean + B5E.mean + B5N.mean + B5O.mean + MD.mean +
  TD.mean + P.mean + E.mean + F.mean + Trip_Period + Day_Type +
  Trait_Anxiety.mean + B5C.mean + Gender + (1|P_ID),
  data = Df_rlevel, REML=FALSE)
```

## Full model summary

```
full_model_summary
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula: HR_Raw.mean ~ 1 + Weather.mean + JF.mean + Speed.mean + ATP.mean +
## Accel_Energy.mean + Rot_Energy.mean + State_Anxiety.mean +
## B5A.mean + B5E.mean + B5N.mean + B5O.mean + MD.mean + TD.mean +
## P.mean + E.mean + F.mean + Trip_Period + Day_Type + Trait_Anxiety.mean +
## B5C.mean + Gender + (1 | P_ID)
## Data: Df_rlevel
##
##      AIC      BIC   logLik deviance df.resid
##  1848.7   1934.0   -900.3   1800.7      235
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.6004 -0.6016 -0.1016  0.5421  3.3624
##
## Random effects:
##  Groups   Name                Variance Std.Dev.
##  P_ID      (Intercept)         9.27     3.045
##  Residual                    56.14     7.493
## Number of obs: 259, groups: P_ID, 21
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   38.74356   13.55569   16.01830   2.858 0.011379 *
## Weather.mean  -1.02415    0.93699   254.36416  -1.093 0.275419
## JF.mean        0.18999    0.87211   237.08868   0.218 0.827730
## Speed.mean     0.08687    0.05754   212.93050   1.510 0.132578
## ATP.mean       0.02036    0.06938   16.60734   0.293 0.772866
## Accel_Energy.mean -1.61783    2.16228   251.24941  -0.748 0.455034
## Rot_Energy.mean -1.30256    2.05338   254.70745  -0.634 0.526424
## State_Anxiety.mean -0.02910    0.07915   258.98506  -0.368 0.713445
## B5A.mean       0.38645    0.66963   17.71547   0.577 0.571126
## B5E.mean      -0.02505    0.44873   11.73578  -0.056 0.956416
## B5N.mean      -1.06732    0.78059   13.08628  -1.367 0.194548
## B5O.mean      -0.47541    0.67838   16.19923  -0.701 0.493373
## MD.mean       -0.30827    0.67176   257.83519  -0.459 0.646689
## TD.mean       -0.28388    0.49201   255.00348  -0.577 0.564464
## P.mean         0.06814    0.58521   208.54467   0.116 0.907417
## E.mean        -0.05191    0.75164   240.82192  -0.069 0.945001
```

```
## F.mean          -0.48826    0.52460 256.15070 -0.931 0.352867
## Trip_PeriodAfternoon  4.49155    1.05852 244.97310  4.243 3.13e-05 ***
## Day_TypeWeekDays    2.72007    1.29177 241.31635  2.106 0.036265 *
## Trait_Anxiety.mean   0.68214    0.17630  15.65521  3.869 0.001409 **
## B5C.mean            2.70801    0.74751  17.01653  3.623 0.002100 **
## GenderMale          9.34159    2.26022  14.85128  4.133 0.000902 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Correlation matrix not shown by default, as p = 22 > 12.
## Use print(x, correlation=TRUE) or
##     vcov(x)         if you need it
```

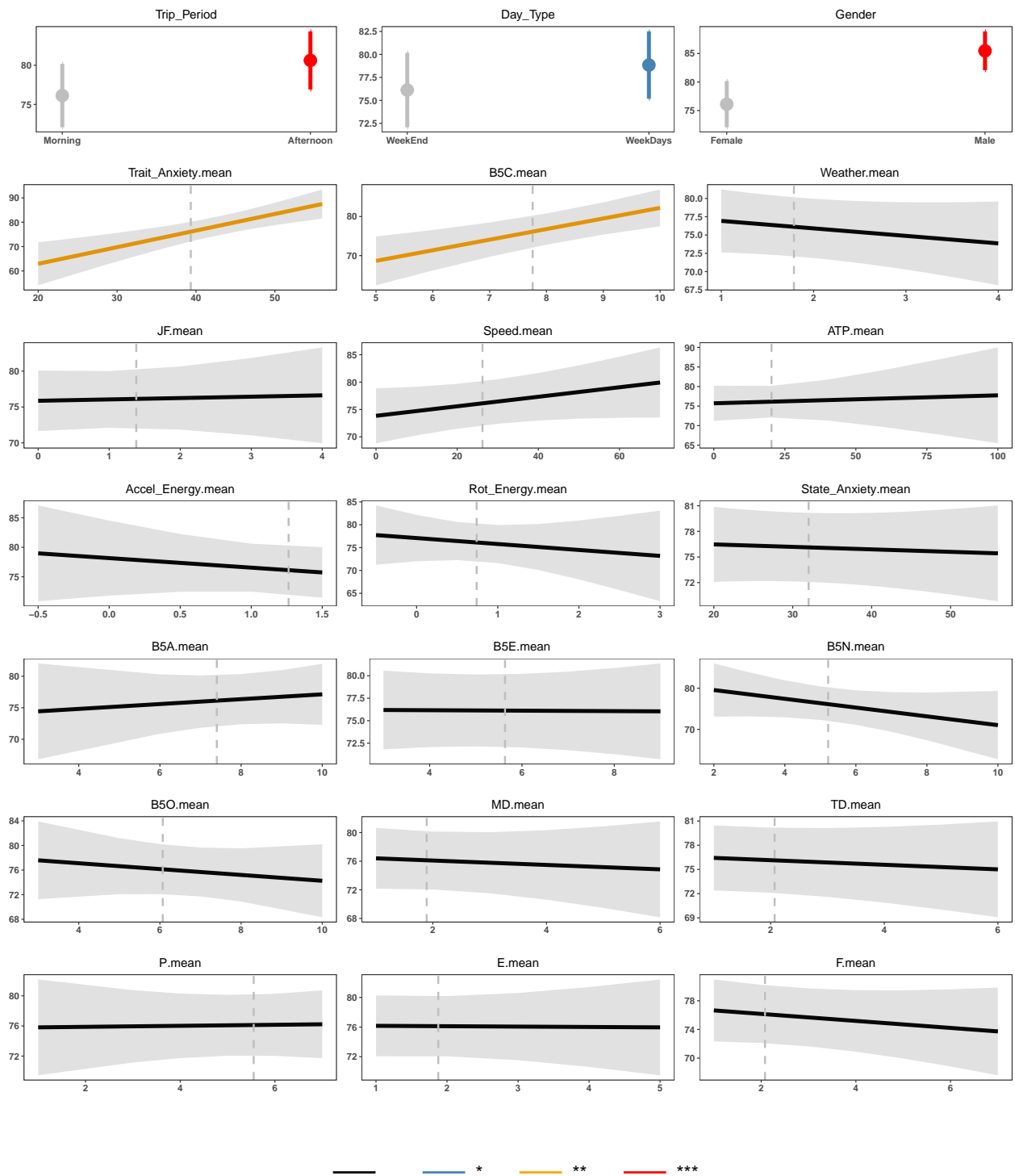
## Full Model AIC

```
aic.full.model
```

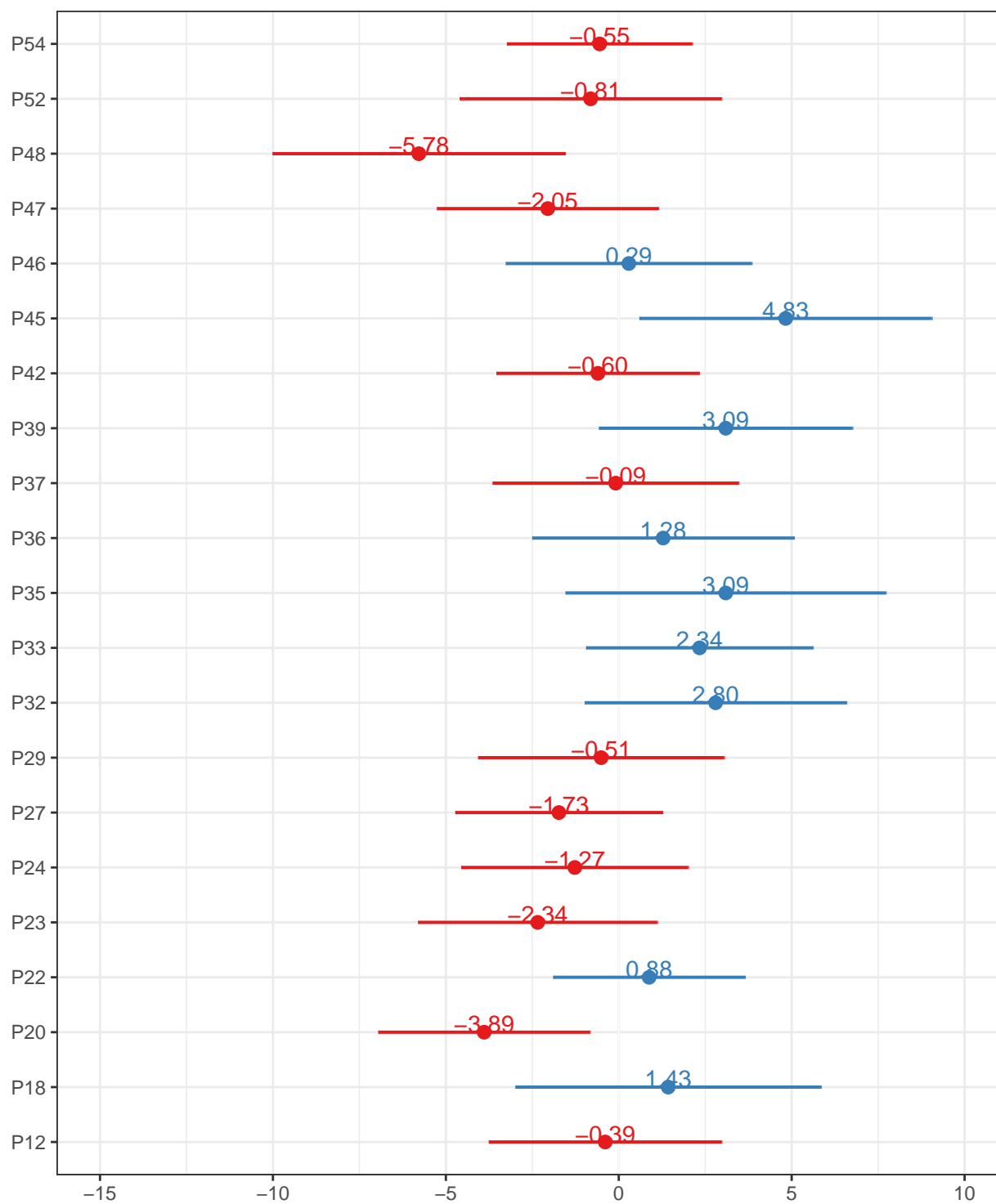
```
## [1] 1848.678
```

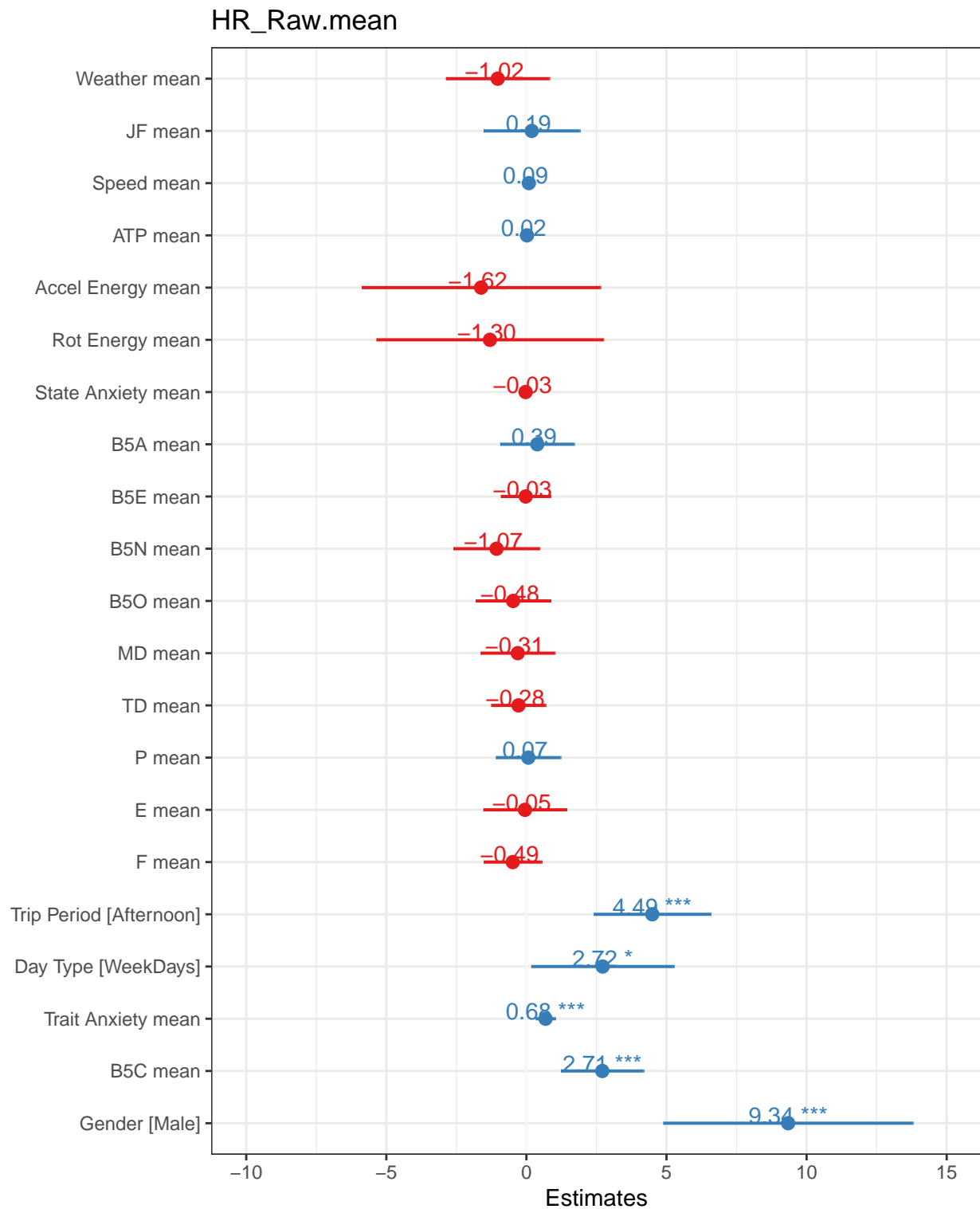
# Full model Plots

## Predictor Plot



# Random effects





## Backward Elimination (BE)

### BE steps

```
step_bw
```

```
## Backward reduced random-effect table:
##
##           Eliminated npar  logLik    AIC    LRT Df Pr(>Chisq)
## <none>                24 -900.34 1848.7
## (1 | P_ID)            0   23 -902.68 1851.4 4.6788 1    0.03054 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Backward reduced fixed-effect table:
## Degrees of freedom method: Satterthwaite
##
##           Eliminated  Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## B5E.mean             1    0.17   0.17     1   11.736  0.0031 0.9564162
## E.mean                2    0.26   0.26     1  241.173  0.0046 0.9458188
## P.mean                3    1.08   1.08     1  213.580  0.0193 0.8896039
## JF.mean               4    2.49   2.49     1  221.130  0.0443 0.8334946
## ATP.mean              5    4.52   4.52     1   18.098  0.0806 0.7796942
## State_Anxiety.mean    6    6.71   6.71     1  254.044  0.1195 0.7298247
## MD.mean               7   13.39  13.39     1  257.307  0.2393 0.6251606
## Rot_Energy.mean       8   25.83  25.83     1  255.000  0.4613 0.4976511
## B5A.mean              9   22.74  22.74     1   19.153  0.4046 0.5322566
## TD.mean              10   38.22  38.22     1  257.040  0.6814 0.4098609
## Weather.mean          11   53.19  53.19     1  255.796  0.9445 0.3320477
## B50.mean              12   83.83  83.83     1   17.205  1.4847 0.2394843
## B5N.mean              13   43.73  43.73     1   18.340  0.7776 0.3892787
## Accel_Energy.mean     14   65.83  65.83     1  240.503  1.1741 0.2796371
## F.mean                15  104.83 104.83     1  258.930  1.8657 0.1731500
## Speed.mean            0  312.97 312.97     1  204.871  5.5263 0.0196826
## Trip_Period           0 1079.86 1079.86     1  246.839 19.0675 1.858e-05
## Day_Type              0  294.81 294.81     1  249.058  5.2056 0.0233588
## Trait_Anxiety.mean    0 1019.44 1019.44     1   16.557 18.0007 0.0005787
## B5C.mean              0   695.23 695.23     1   17.136 12.2760 0.0026950
## Gender                0 1098.03 1098.03     1   17.116 19.3885 0.0003827
##
## B5E.mean
## E.mean
## P.mean
## JF.mean
## ATP.mean
## State_Anxiety.mean
## MD.mean
## Rot_Energy.mean
## B5A.mean
## TD.mean
## Weather.mean
## B50.mean
## B5N.mean
## Accel_Energy.mean
```



```
## F.mean
## Speed.mean      *
## Trip_Period     ***
## Day_Type        *
## Trait_Anxiety.mean ***
## B5C.mean        **
## Gender          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Model found:
## HR_Raw.mean ~ Speed.mean + Trip_Period + Day_Type + Trait_Anxiety.mean + B5C.mean + Gender + (1 | P_ID)
```

## BE summary

### suggested\_bw\_model\_summary

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula:
## HR_Raw.mean ~ Speed.mean + Trip_Period + Day_Type + Trait_Anxiety.mean +
## B5C.mean + Gender + (1 | P_ID)
## Data: Df_rlevel
##
##      AIC      BIC   logLik deviance df.resid
##  1826.8   1858.8   -904.4   1808.8     250
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.2352 -0.5903 -0.1093  0.4775  3.3613
##
## Random effects:
## Groups   Name      Variance Std.Dev.
## P_ID      (Intercept) 14.04    3.747
## Residual                56.63    7.526
## Number of obs: 259, groups: P_ID, 21
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   35.11504    8.60607   20.11570   4.080 0.000577 ***
## Speed.mean     0.11290    0.04802  204.87122   2.351 0.019683 *
## Trip_PeriodAfternoon 4.44543    1.01804  246.83914   4.367 1.86e-05 ***
## Day_TypeWeekDays 2.74979    1.20522  249.05779   2.282 0.023359 *
## Trait_Anxiety.mean 0.47565    0.11211   16.55702   4.243 0.000579 ***
## B5C.mean       2.45978    0.70205   17.13610   3.504 0.002695 **
## GenderMale     9.87552    2.24279   17.11572   4.403 0.000383 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Spd.mn Trp_PA Dy_TWD Trt_A. B5C.mn
## Speed.mean   -0.269
## Trp_Prdafter -0.057  0.075
## Dy_TypWkdys  -0.115 -0.005  0.121
```

```
## Trt_Anxtty.m -0.715  0.081 -0.039 -0.024
## B5C.mean      -0.778  0.128 -0.032  0.037  0.203
## GenderMale    -0.458 -0.090  0.008 -0.083  0.401  0.171
```

## BE AIC

```
aic.backward
```

```
## [1] 1826.78
```

## Forward Elimination (FE)

### FE steps

```
step_fw
```

```
## Backward reduced random-effect table:
##
##           Eliminated npar  logLik    AIC    LRT Df Pr(>Chisq)
## <none>                3 -924.80 1855.6
## (1 | P_ID)            0   2 -969.51 1943.0 89.417  1 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Backward reduced fixed-effect table:
## Degrees of freedom method: Satterthwaite
##
##           Eliminated Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
##
## Model found:
## HR_Raw.mean ~ 1 + (1 | P_ID)
```

### FE summary

```
suggested_fw_model_summary
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: HR_Raw.mean ~ 1 + (1 | P_ID)
## Data: Df_rlevel
##
## REML criterion at convergence: 1849.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.1682 -0.6483 -0.1784  0.5171  3.2979
##
## Random effects:
## Groups   Name                Variance Std.Dev.
## P_ID     (Intercept)  50.32      7.093
## Residual                    62.21      7.887
## Number of obs: 259, groups: P_ID, 21
##
## Fixed effects:
##              Estimate Std. Error    df t value Pr(>|t|)
## (Intercept)   87.201      1.639 19.853   53.21 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

### FE AIC

```
aic.forward
```

```
## [1] 1855.595
```

## Both Direction Elimination (BDE)

### BDE steps

```
step_bd
```

```
## Backward reduced random-effect table:
##
##           Eliminated npar  logLik    AIC    LRT Df Pr(>Chisq)
## <none>                3 -924.80 1855.6
## (1 | P_ID)            0   2 -969.51 1943.0 89.417  1 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Backward reduced fixed-effect table:
## Degrees of freedom method: Satterthwaite
##
##           Eliminated Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
##
## Model found:
## HR_Raw.mean ~ 1 + (1 | P_ID)
```

### BE summary

```
suggested_bd_model_summary
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: HR_Raw.mean ~ 1 + (1 | P_ID)
## Data: Df_rlevel
##
## REML criterion at convergence: 1849.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.1682 -0.6483 -0.1784  0.5171  3.2979
##
## Random effects:
## Groups   Name                Variance Std.Dev.
## P_ID     (Intercept)  50.32      7.093
## Residual                    62.21      7.887
## Number of obs: 259, groups: P_ID, 21
##
## Fixed effects:
##              Estimate Std. Error    df t value Pr(>|t|)
## (Intercept)   87.201      1.639 19.853   53.21 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

### BE AIC

```
aic.both
```

```
## [1] 1855.595
```

## Optimal model

- Backward elimination model is suggested one as it has low AIC value compared to Forward elimination and Both Direction elimination.

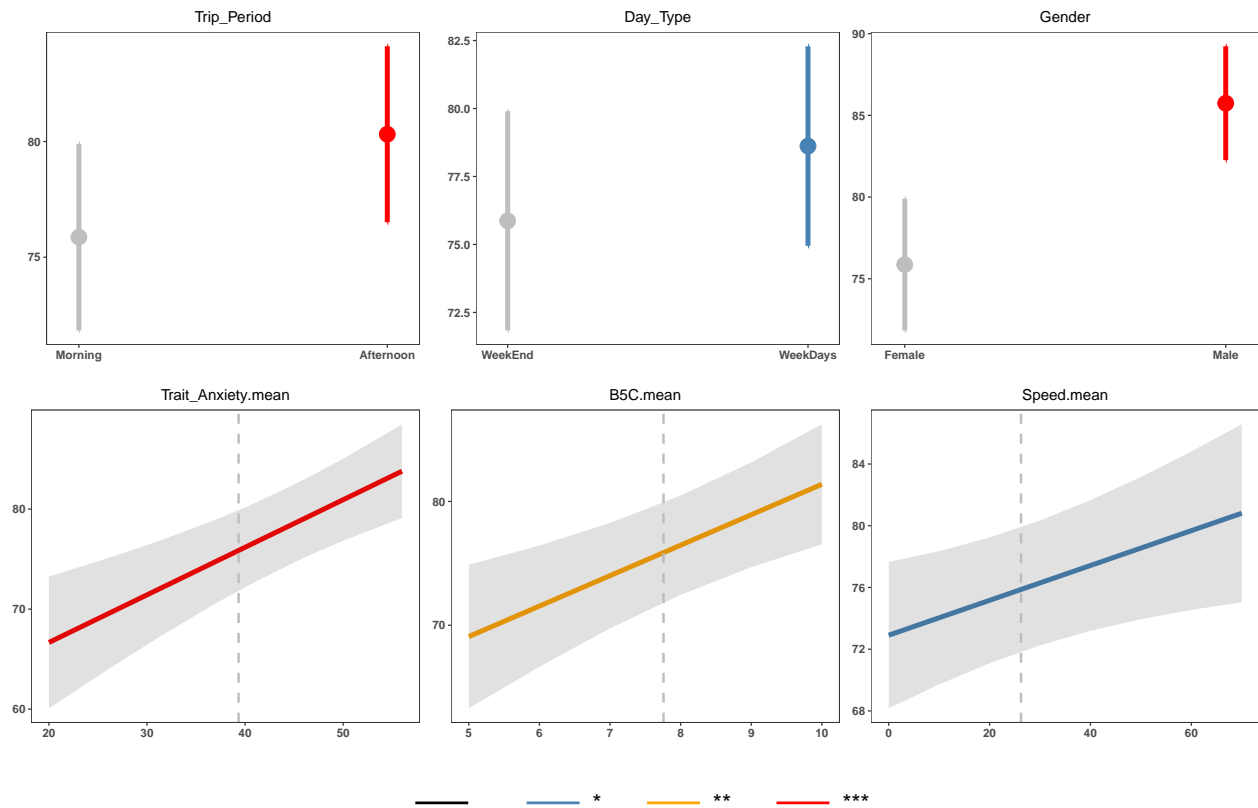
## Optimal model summary

### suggested\_bw\_model\_summary

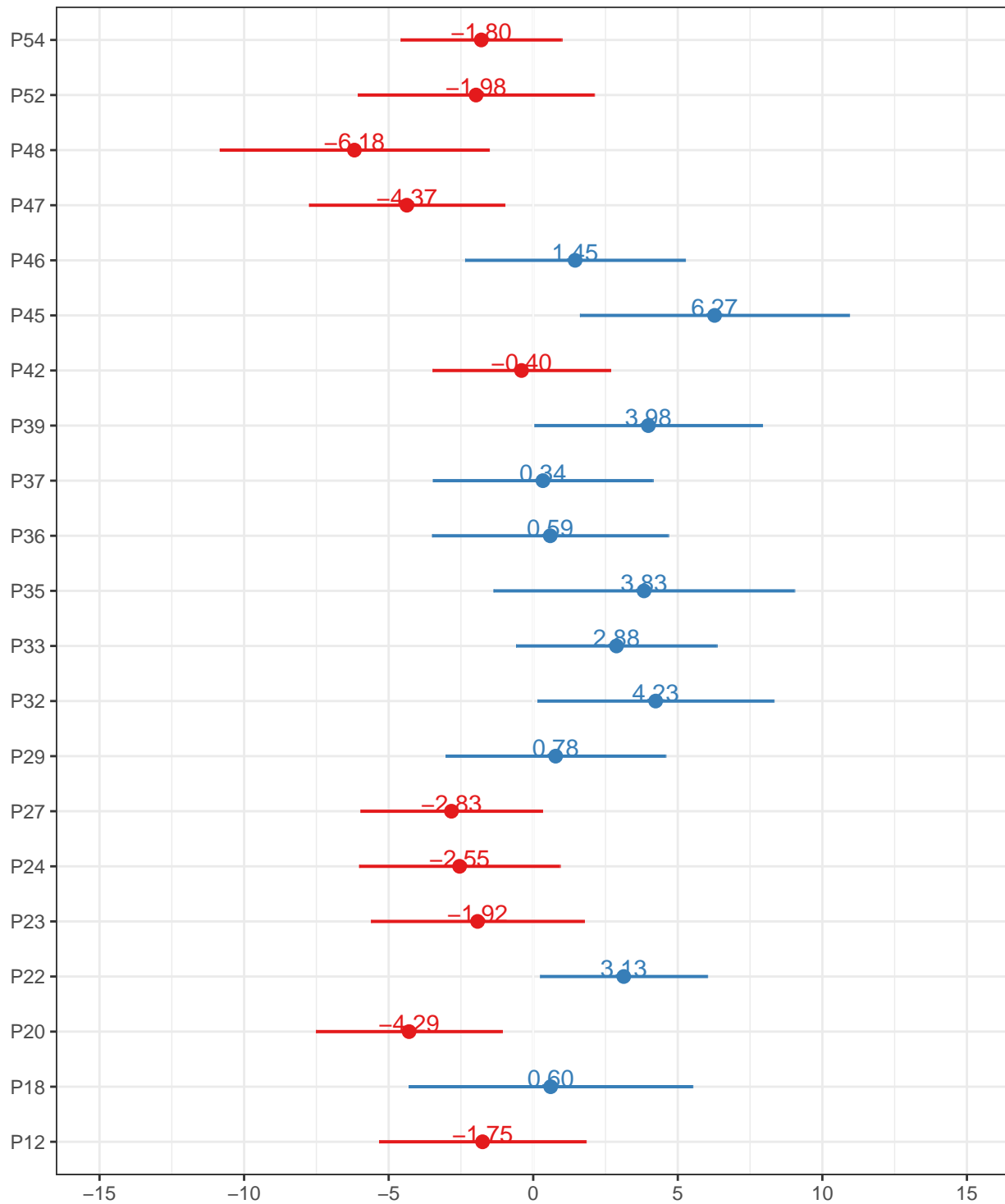
```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula:
## HR_Raw.mean ~ Speed.mean + Trip_Period + Day_Type + Trait_Anxiety.mean +
## B5C.mean + Gender + (1 | P_ID)
## Data: Df_rlevel
##
##          AIC          BIC    logLik deviance df.resid
##    1826.8    1858.8   -904.4   1808.8     250
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.2352 -0.5903 -0.1093  0.4775  3.3613
##
## Random effects:
## Groups   Name            Variance Std.Dev.
## P_ID      (Intercept)  14.04      3.747
## Residual                    56.63      7.526
## Number of obs: 259, groups: P_ID, 21
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)    35.11504    8.60607   20.11570   4.080 0.000577 ***
## Speed.mean       0.11290    0.04802   204.87122   2.351 0.019683 *
## Trip_PeriodAfternoon  4.44543    1.01804   246.83914   4.367 1.86e-05 ***
## Day_TypeWeekDays    2.74979    1.20522   249.05779   2.282 0.023359 *
## Trait_Anxiety.mean  0.47565    0.11211   16.55702   4.243 0.000579 ***
## B5C.mean         2.45978    0.70205   17.13610   3.504 0.002695 **
## GenderMale        9.87552    2.24279   17.11572   4.403 0.000383 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Spd.mn Trp_PA Dy_TWD Trt_A. B5C.mn
## Speed.mean   -0.269
## Trp_PrdAftr  -0.057  0.075
## Dy_TypWkDys  -0.115 -0.005  0.121
## Trt_Anxtty.m -0.715  0.081 -0.039 -0.024
## B5C.mean     -0.778  0.128 -0.032  0.037  0.203
## GenderMale   -0.458 -0.090  0.008 -0.083  0.401  0.171
```

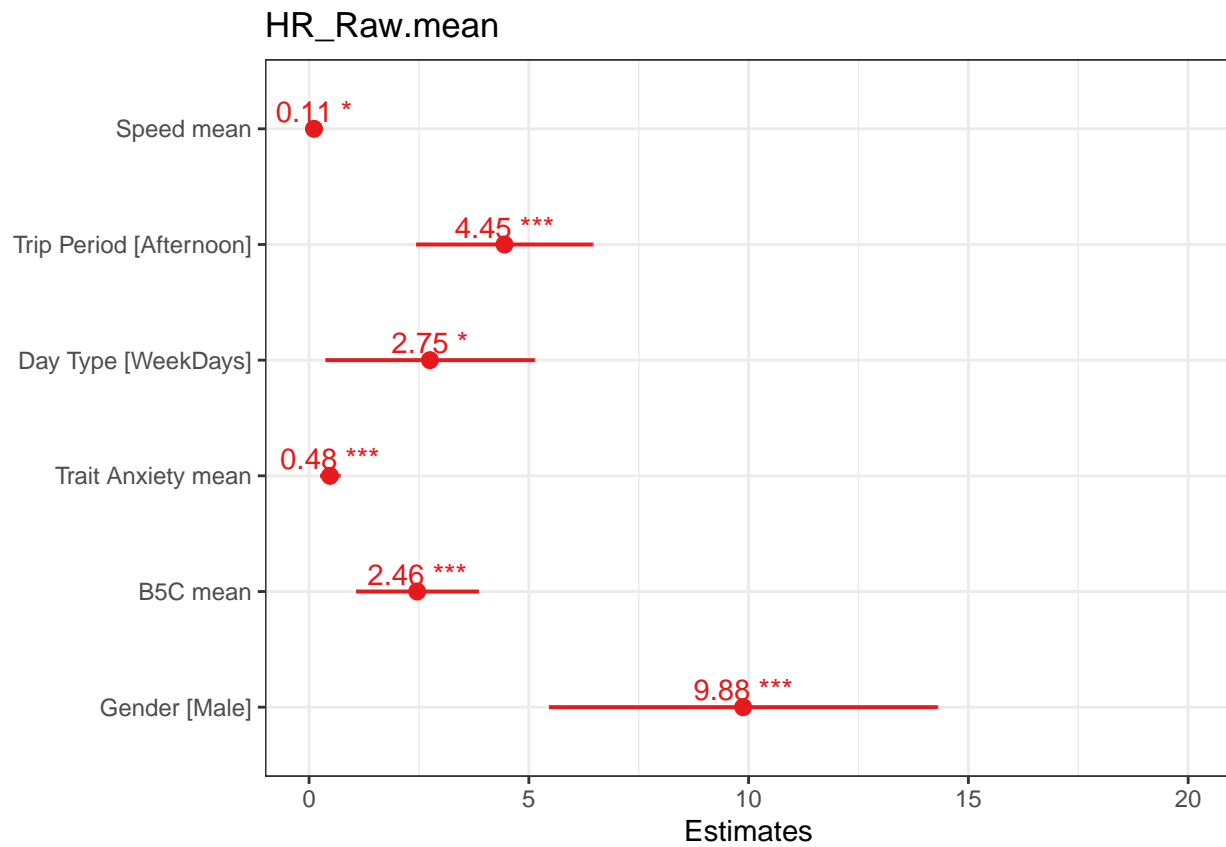
## Optimal model plots

### Predictor plot



# Random effects





## Conclusion

- The final predictors that effect the HR\_Raw are Speed.mean, Trip\_PeriodAfternoon, Day\_TypeWeekDays, Trait\_Anxiety.mean, B5C.mean and GenderMale.