Preliminar-analysis

Gabriel Robaina

2023-11-10

Preliminar analysis

Load the CSV data in the AOV to test 3 factors with 2 levels each: * Time of day: BUSINESS_HOUR and OFF_HOUR * Week Period: WEEKDAY and WEEKEND * Serverless provider: Lambda and GCF

```
data.aov = aov(LATENCY_SECONDS ~ SYSTEM_NAME * TIME_OF_DAY * WEEK_PERIOD * OPERATION_TYPE,
               data=data)
summary(data.aov)
```

```
##                                          Df Sum Sq Mean Sq  F value   Pr(>F)
## SYSTEM_NAME                               1  0.927   0.927  223.481  < 2e-16 ***
## TIME_OF_DAY                               1  0.052   0.052   12.600 0.000705 ***
## WEEK_PERIOD                               1  0.001   0.001    0.195 0.659818
## OPERATION_TYPE                            1 12.416  12.416 2992.404  < 2e-16 ***
## SYSTEM_NAME:TIME_OF_DAY                   1  0.008   0.008    1.871 0.175914
## SYSTEM_NAME:WEEK_PERIOD                   1  0.001   0.001    0.202 0.654523
## SYSTEM_NAME:OPERATION_TYPE                1  0.987   0.987  237.824  < 2e-16 ***
## TIME_OF_DAY:OPERATION_TYPE                1  0.006   0.006    1.485 0.227169
## WEEK_PERIOD:OPERATION_TYPE                1  0.000   0.000    0.022 0.882562
## SYSTEM_NAME:TIME_OF_DAY:OPERATION_TYPE    1  0.009   0.009    2.229 0.140067
## SYSTEM_NAME:WEEK_PERIOD:OPERATION_TYPE    1  0.000   0.000    0.105 0.747158
## Residuals                                68  0.282   0.004
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Allocation of variation

```
SS = anova(data.aov)["Sum Sq"]
SST = sum(SS)
round(100*SS/sum(SS), 2)
```

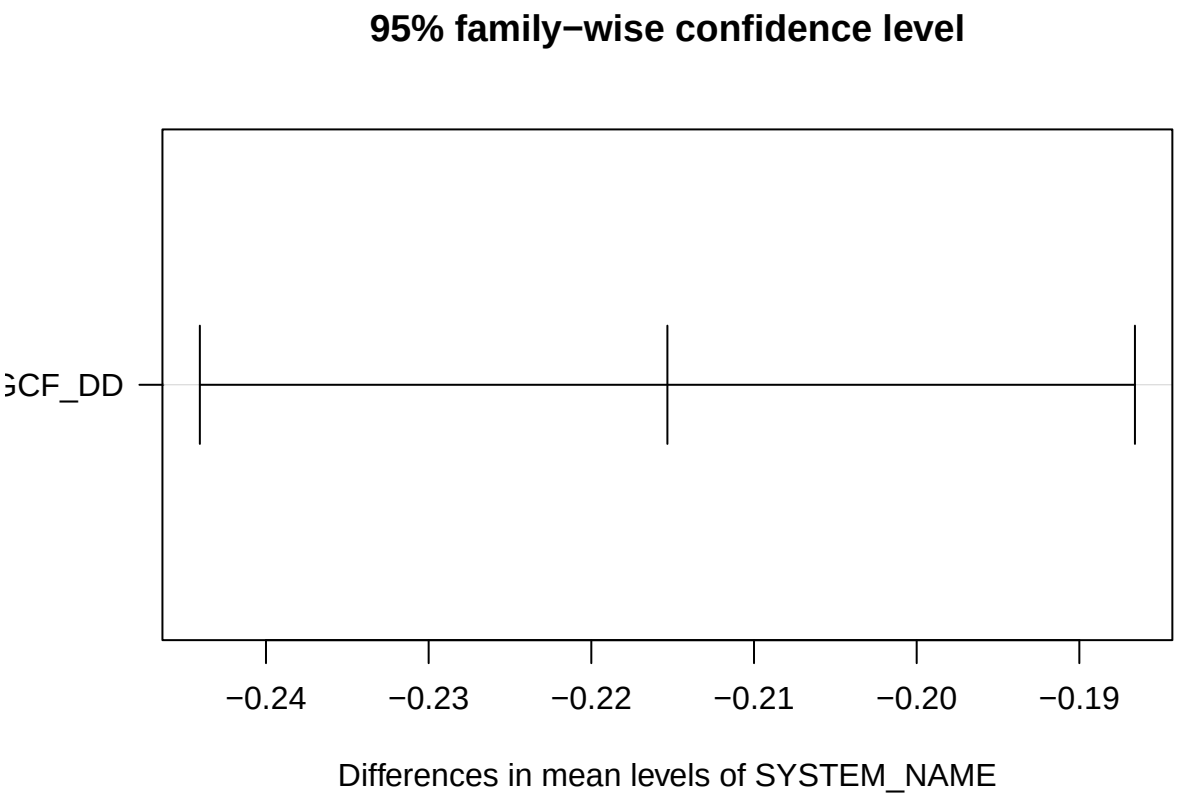
```
##                                          Sum Sq
## SYSTEM_NAME                                6.31
## TIME_OF_DAY                                0.36
## WEEK_PERIOD                                0.01
## OPERATION_TYPE                            84.52
## SYSTEM_NAME:TIME_OF_DAY                    0.05
## SYSTEM_NAME:WEEK_PERIOD                    0.01
```

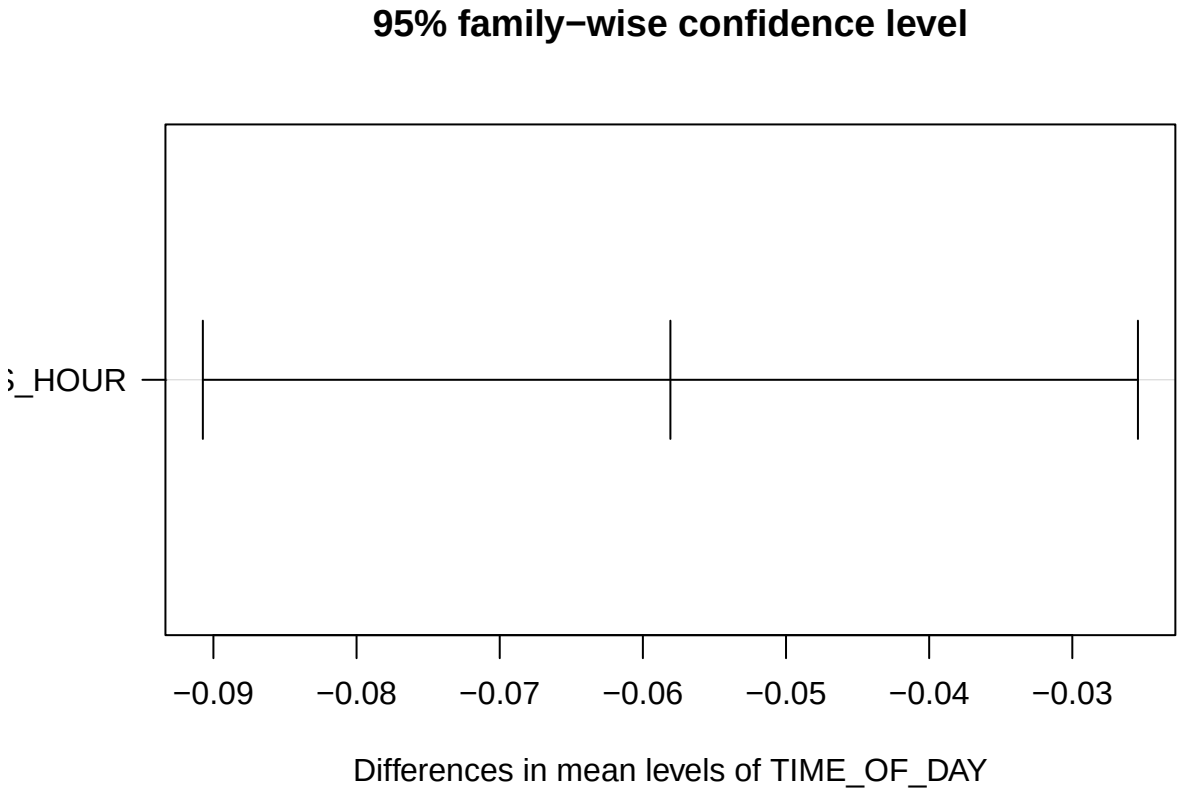
```
## SYSTEM_NAME:OPERATION_TYPE                6.72
## TIME_OF_DAY:OPERATION_TYPE                0.04
## WEEK_PERIOD:OPERATION_TYPE                0.00
## SYSTEM_NAME:TIME_OF_DAY:OPERATION_TYPE    0.06
## SYSTEM_NAME:WEEK_PERIOD:OPERATION_TYPE    0.00
## Residuals                                 1.92
```

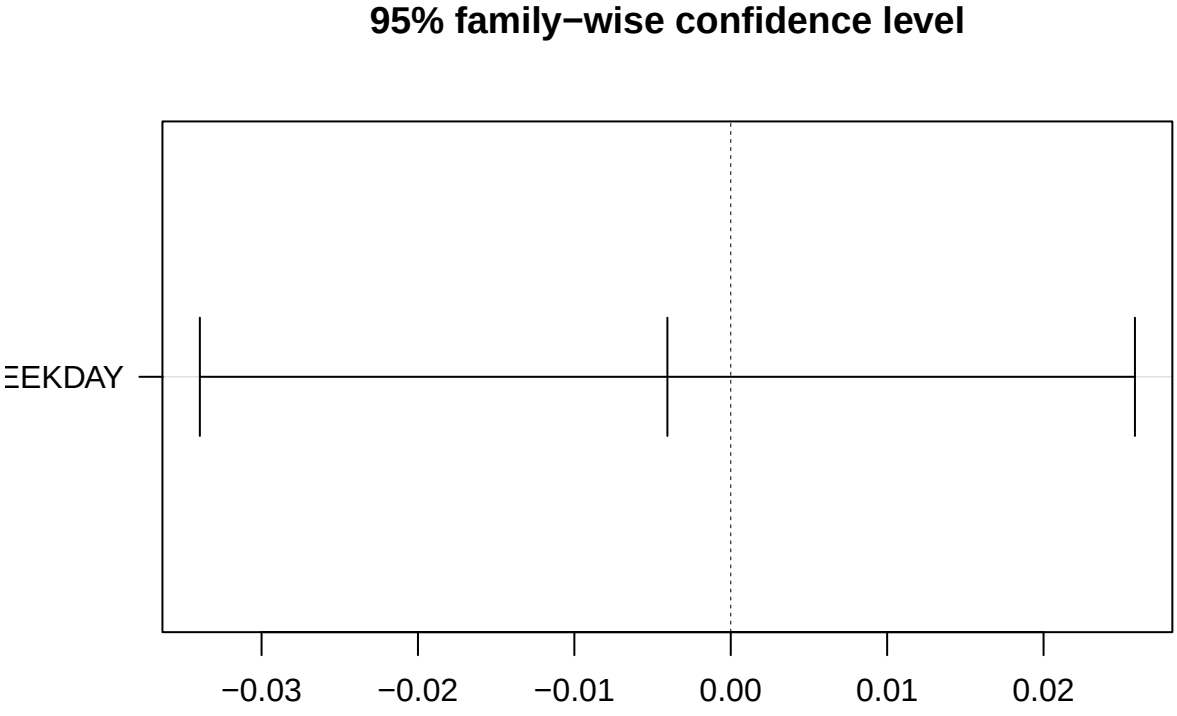
In this analysis most of the variance is coming from the serverless platform, operation type and their interaction.

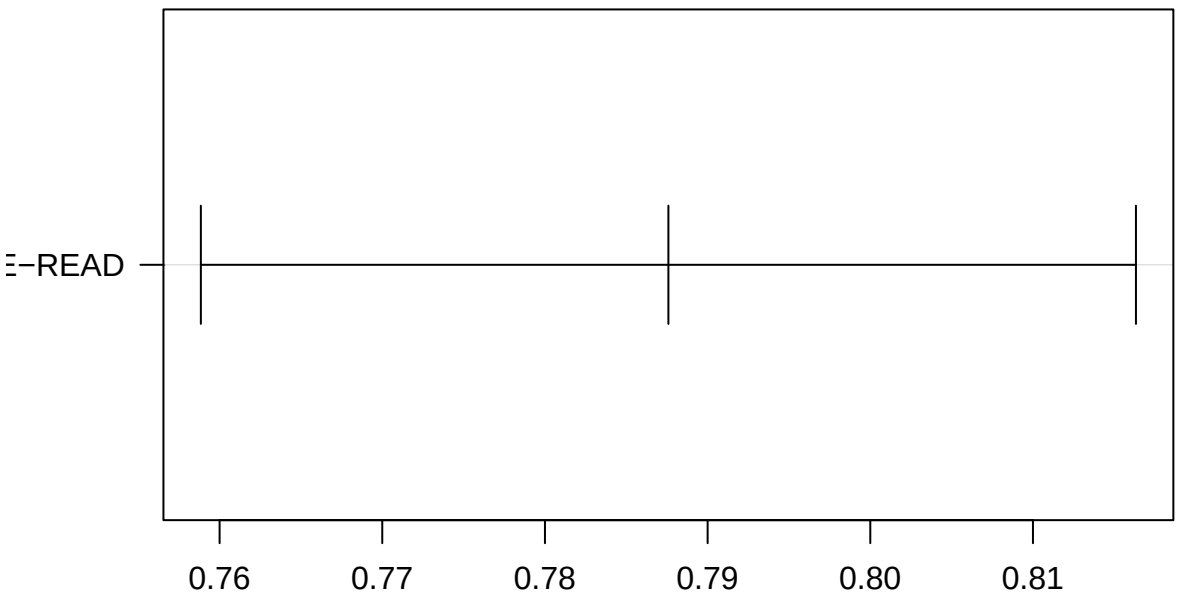
Tukey test

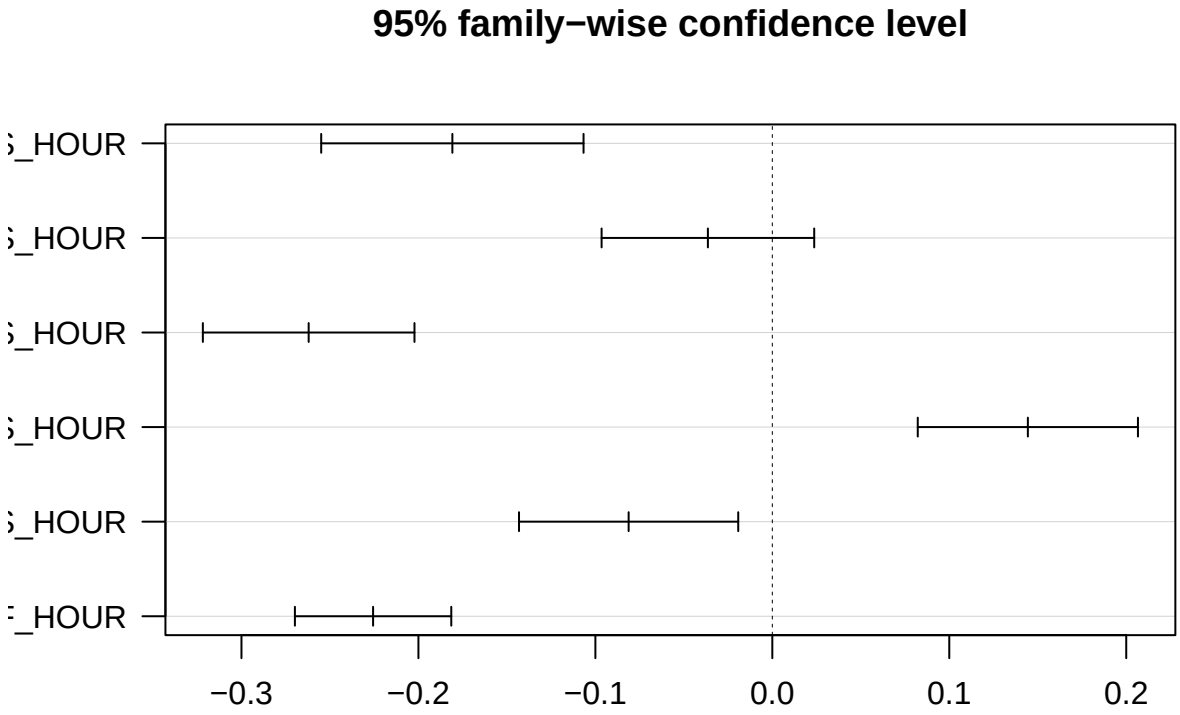
```
plot(tukey,las=1,tcl = -.6)
```

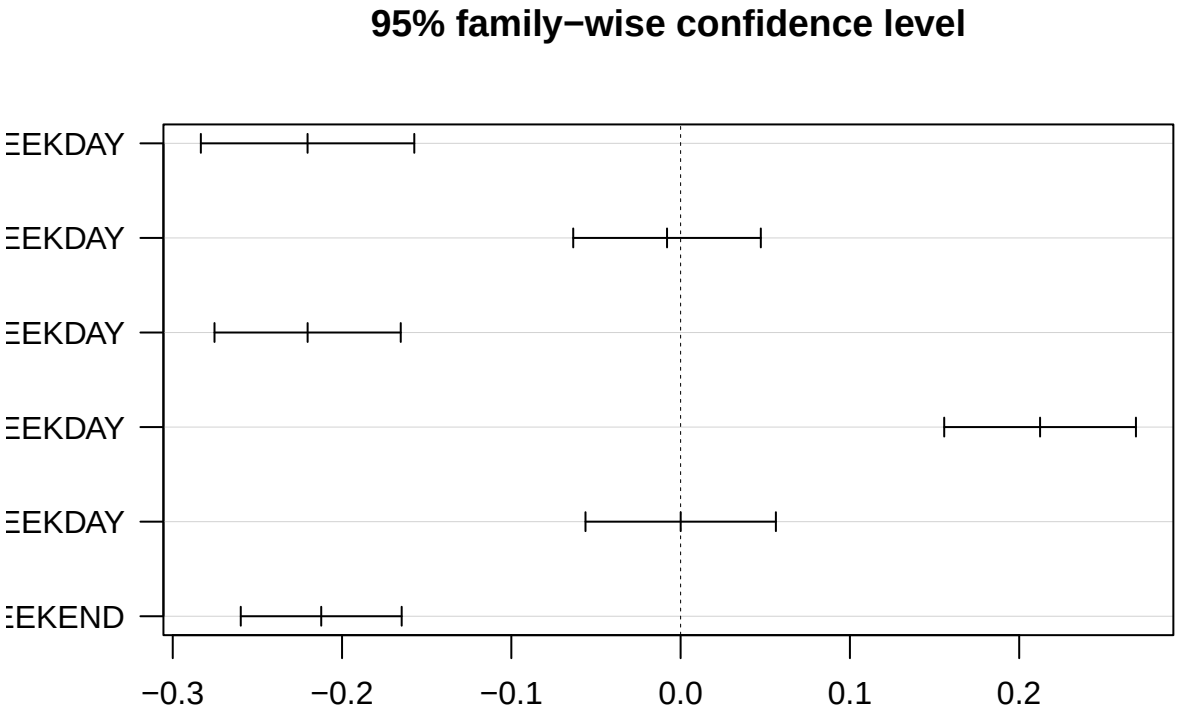


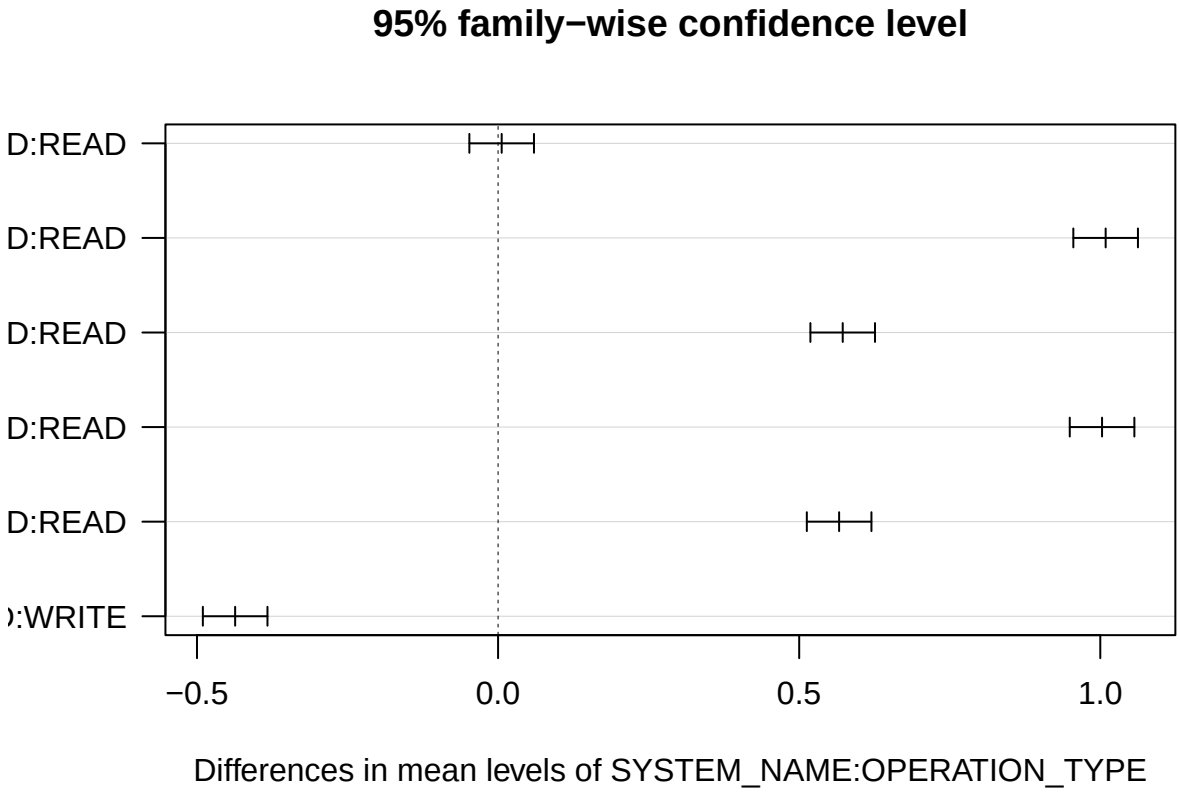

95% family−wise confidence level
S_HOUR
−0.09
−0.08
−0.07
−0.06
−0.05
−0.04
−0.03
Differences in mean levels of TIME_OF_DAY

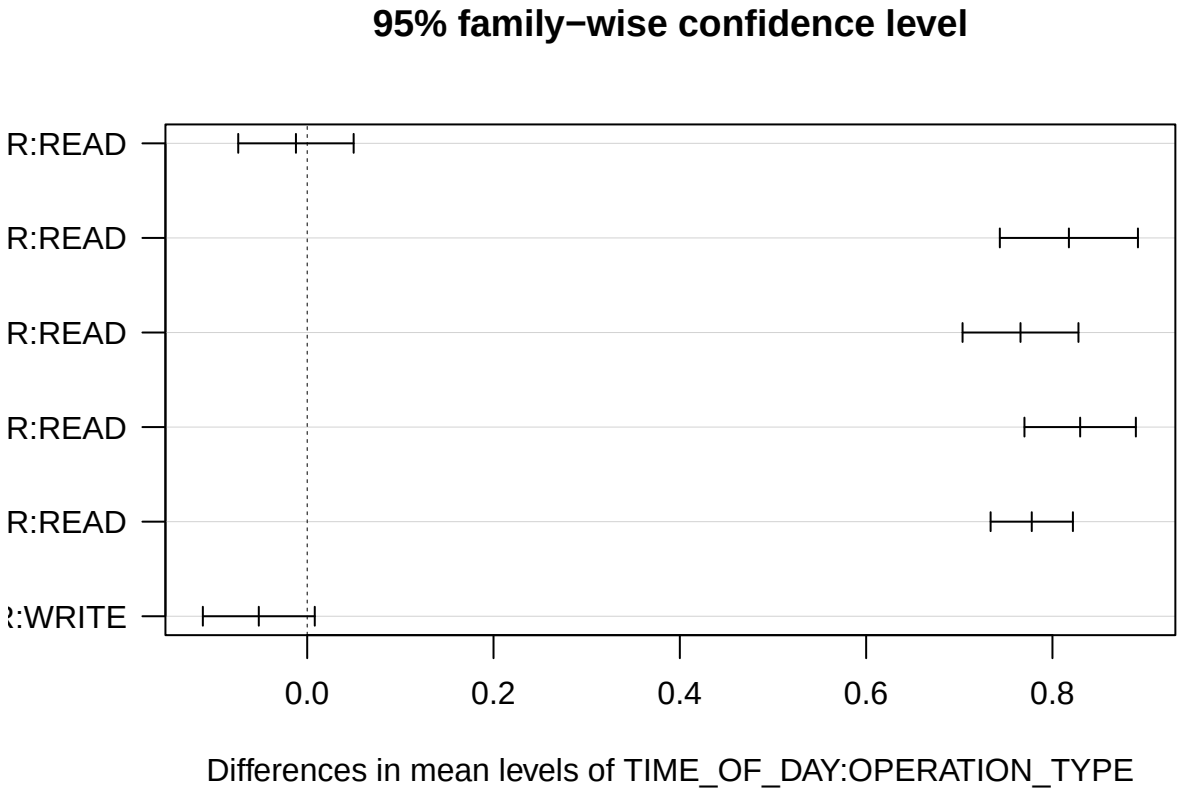

95% family−wise confidence level
ΞEKDAY
−0.03
−0.02
−0.01
0.00
0.01
0.02
Differences in mean levels of WEEK_PERIOD

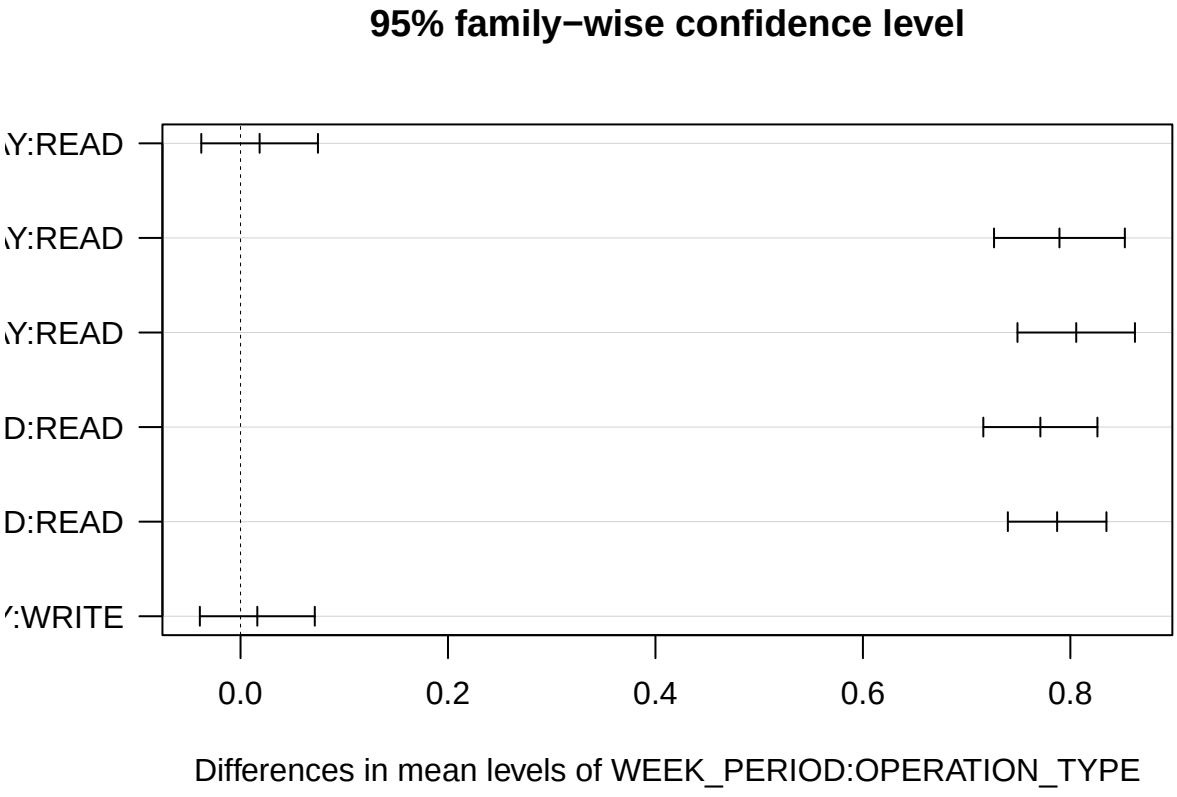

95% family−wise confidence level
E−READ
0.76
0.77
0.78
0.79
0.80
0.81
Differences in mean levels of OPERATION_TYPE

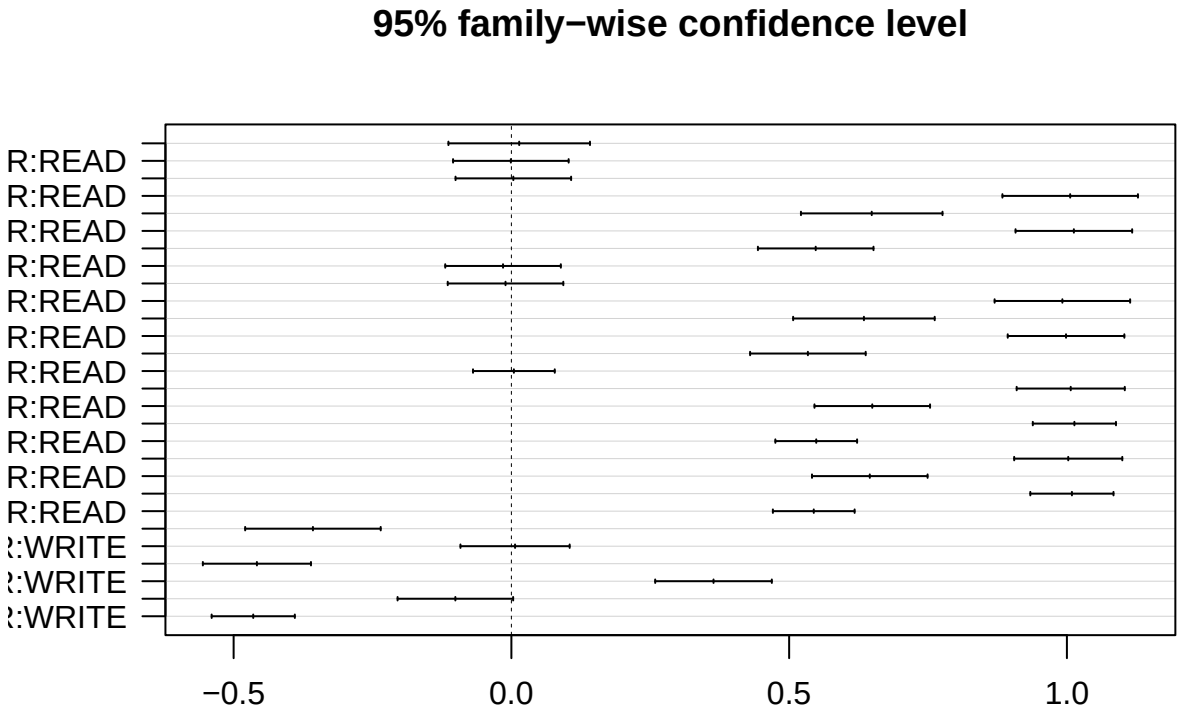

95% family−wise confidence level
S_HOUR
S_HOUR
S_HOUR
S_HOUR
S_HOUR
F_HOUR
−0.3
−0.2
−0.1
0.0
0.1
0.2
Differences in mean levels of SYSTEM_NAME:TIME_OF_DAY

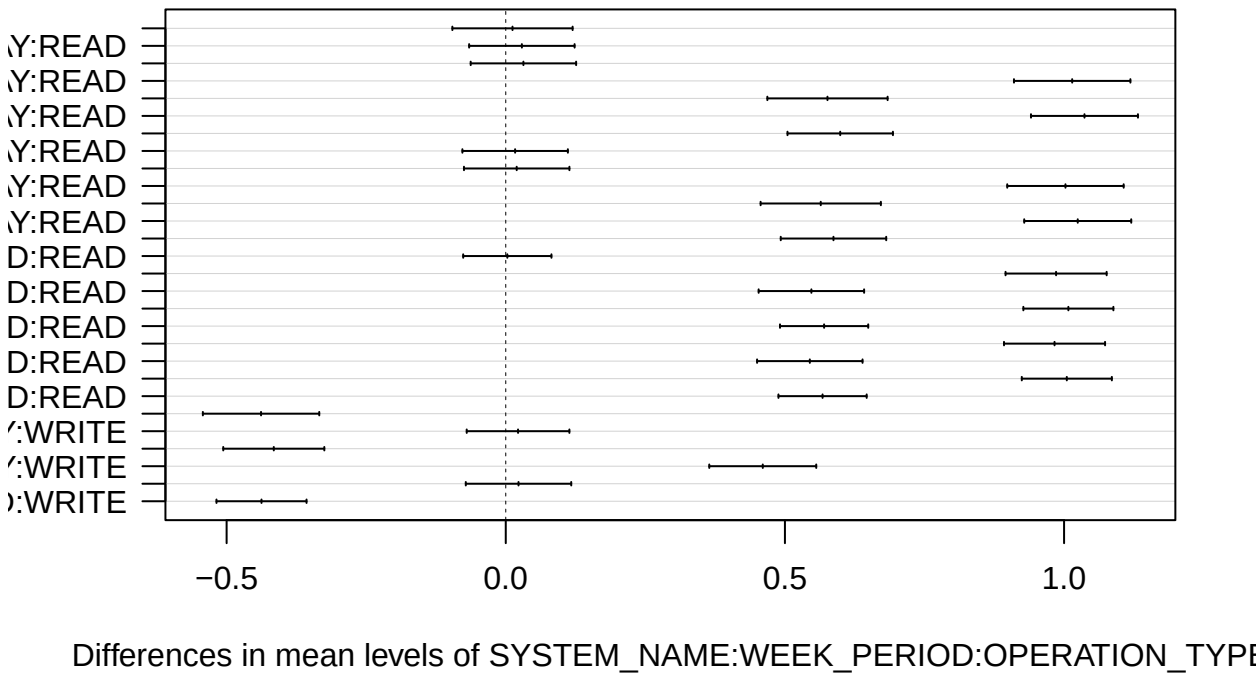

95% family−wise confidence level
ΞEKDAY
ΞEKDAY
ΞEKDAY
ΞEKDAY
ΞEKDAY
ΞEKEND
−0.3
−0.2
−0.1
0.0
0.1
0.2
Differences in mean levels of SYSTEM_NAME:WEEK_PERIOD


95% family−wise confidence level
D:READ
D:READ
D:READ
D:READ
D:READ
):WRITE
−0.5
0.0
0.5
1.0
Differences in mean levels of SYSTEM_NAME:OPERATION_TYPE


95% family−wise confidence level
R:READ
R:READ
R:READ
R:READ
R:READ
:WRITE
0.0
0.2
0.4
0.6
0.8
Differences in mean levels of TIME_OF_DAY:OPERATION_TYPE


95% family−wise confidence level
AY:READ
AY:READ
AY:READ
D:READ
D:READ
':WRITE
0.0
0.2
0.4
0.6
0.8
Differences in mean levels of WEEK_PERIOD:OPERATION_TYPE


95% family−wise confidence level
R:READ
R:READ
R:READ
R:READ
R:READ
R:READ
R:READ
R:READ
R:READ
R:READ
R:READ
:WRITE
:WRITE
:WRITE
−0.5
0.0
0.5
1.0
Differences in mean levels of SYSTEM_NAME:TIME_OF_DAY:OPERATION_TYPE



1. The confidence interval for `SYSTEM_NAME` is negative for `LAMBDA_DD - GCF_DD`. This means that latency time for GCF was higher, hence Lambda is faster.
2. Confidence interval for `OPERATION_TYPE` is positive for `WRITE - READ`. This means that write is higher, hence reads are faster as expected.
3. Lambda reads and GCF reads have no significant differences, so they are the same.
4. Confidence interval for `LAMBDA_DD:WRITE - GCF_DD:WRITE` is negative. This means that writes in GCF are slower than Lambda.
5. Week period is not significant.
6. Time of day is significant for 95% confidence level. Its confidence interval is negative for `OFF-HOUR - BUSINESS_HOUR`. This means that business hours are slower.