

# 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 \* Day of week: WEEKDAY and WEEKEND \* Serverless provider: Lambda and GCF

```
data.aov = aov(LATENCY_SECONDS ~ SYSTEM_NAME * TIME_OF_DAY * DAY_OF_WEEK * 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 ***
## DAY_OF_WEEK                   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:DAY_OF_WEEK       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
## DAY_OF_WEEK: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:DAY_OF_WEEK: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)
```

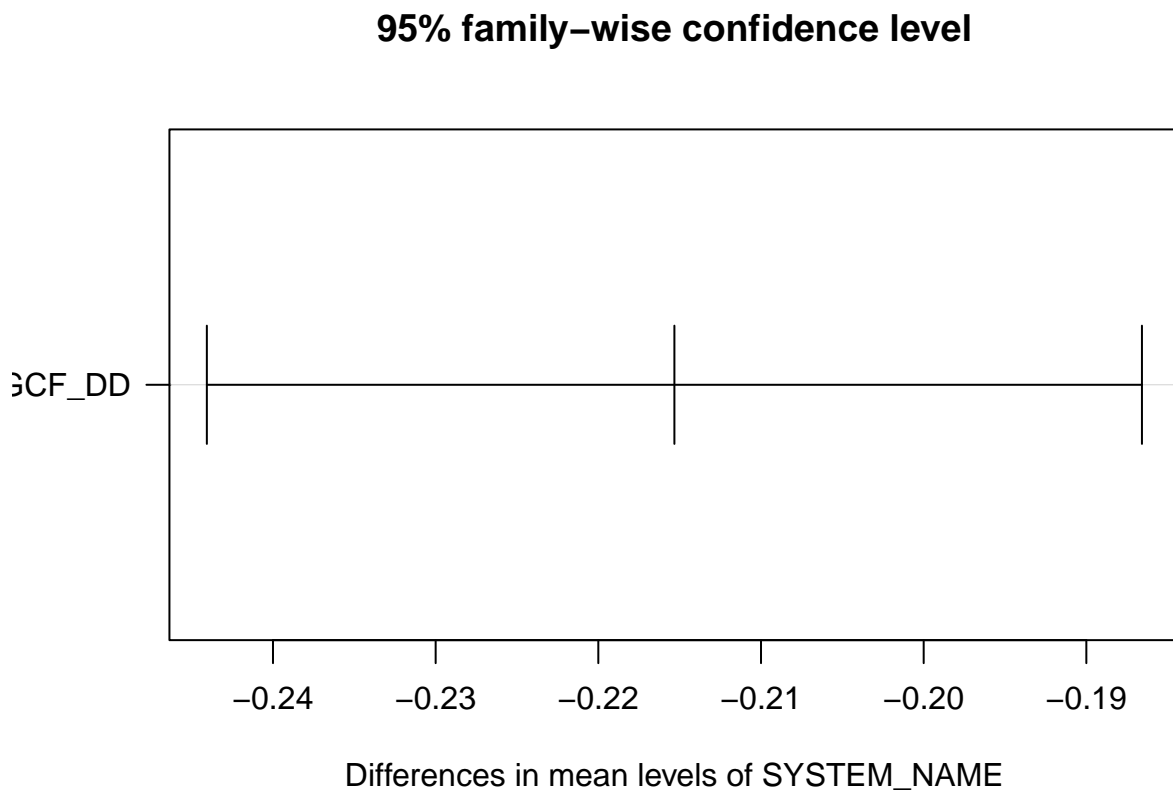
```
##                               Sum Sq
## SYSTEM_NAME                   6.31
## TIME_OF_DAY                   0.36
## DAY_OF_WEEK                   0.01
## OPERATION_TYPE                84.52
## SYSTEM_NAME:TIME_OF_DAY       0.05
## SYSTEM_NAME:DAY_OF_WEEK       0.01
```

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

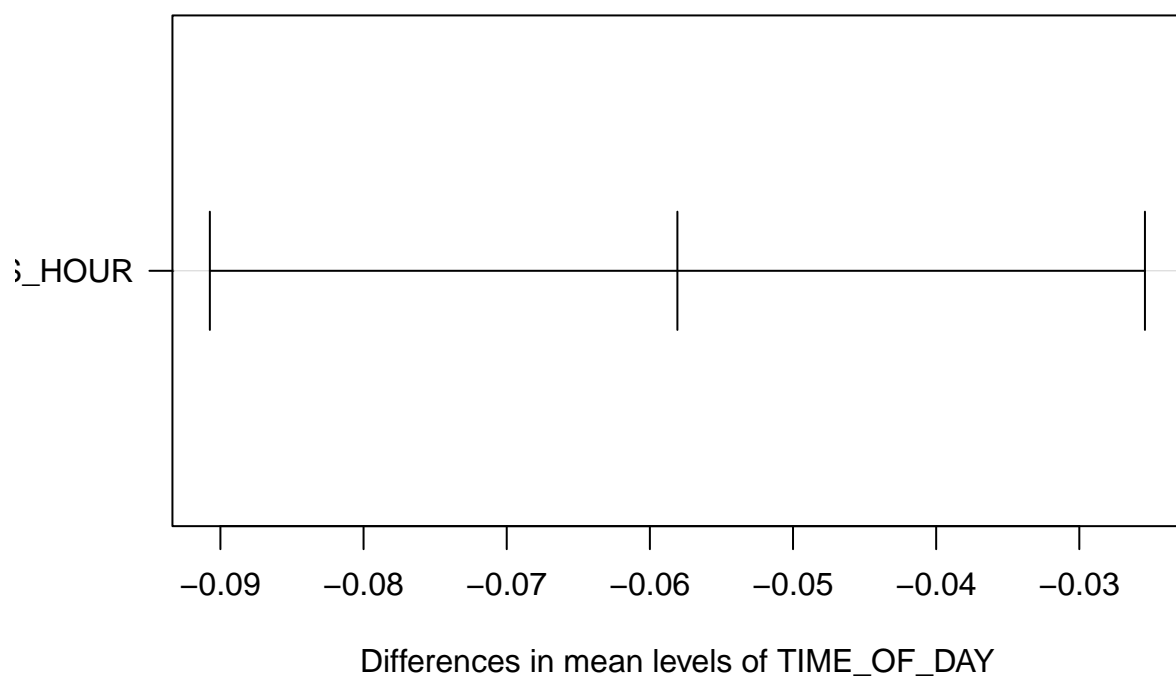
In this analysis most of the variance is coming from the serverless platform, operation type and their interaction.

## Tukey test

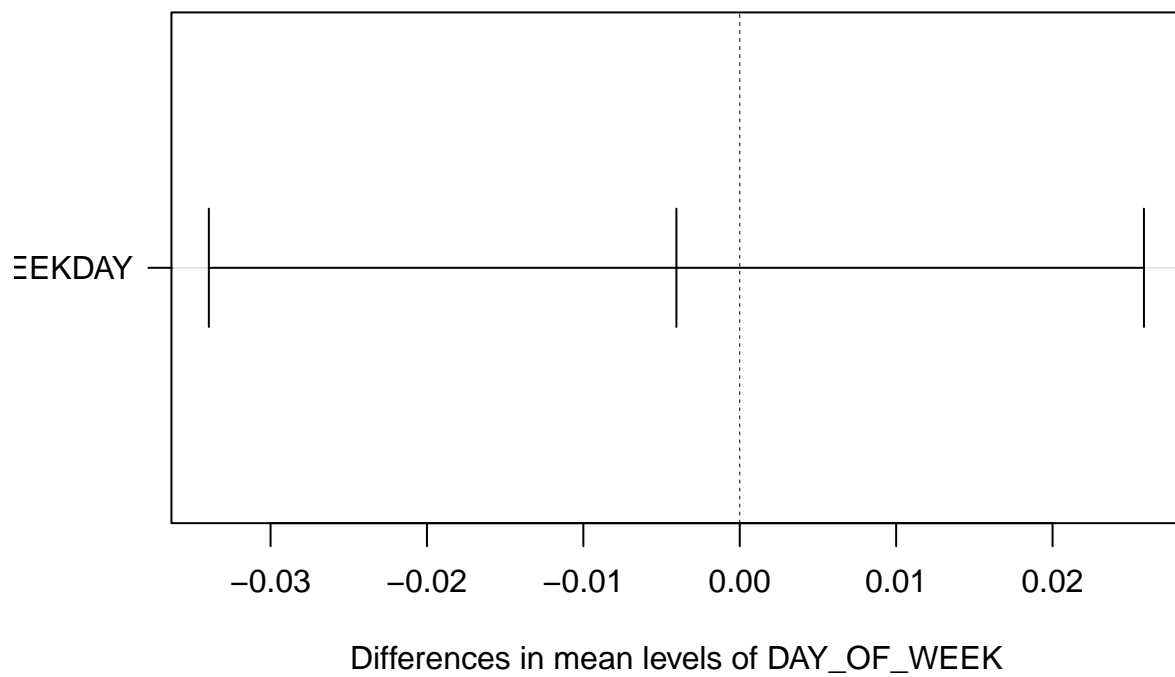
```
plot(tukey,las=1,tcl = -.6)
```



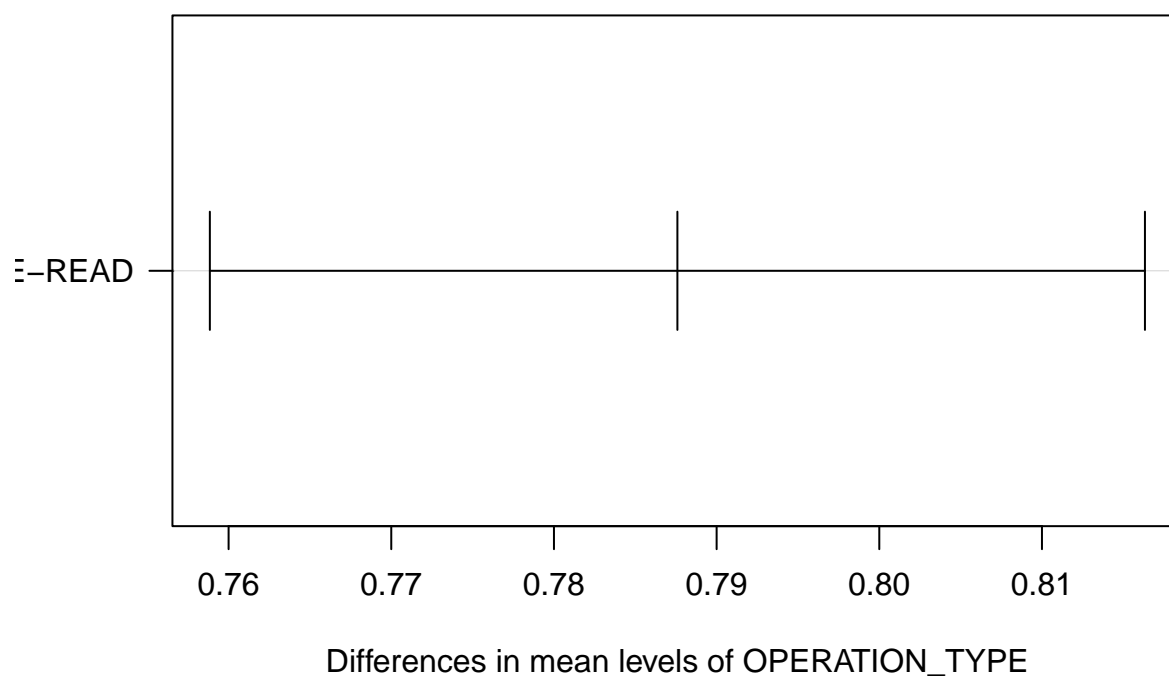
**95% family-wise confidence level**



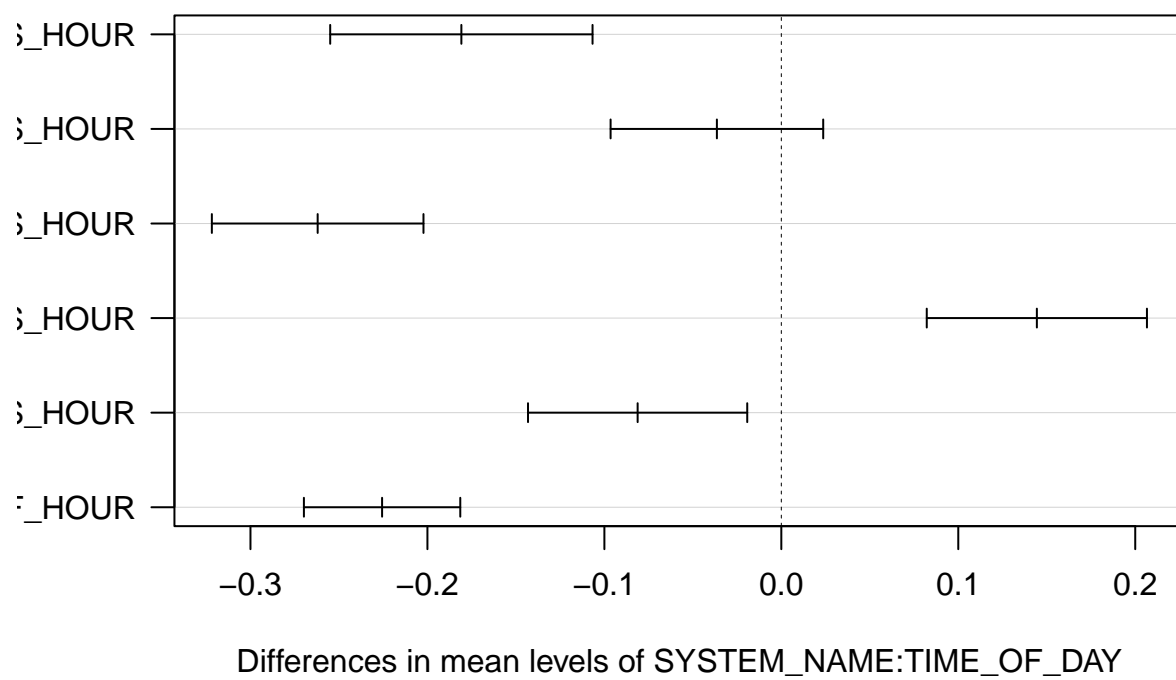
### 95% family-wise confidence level



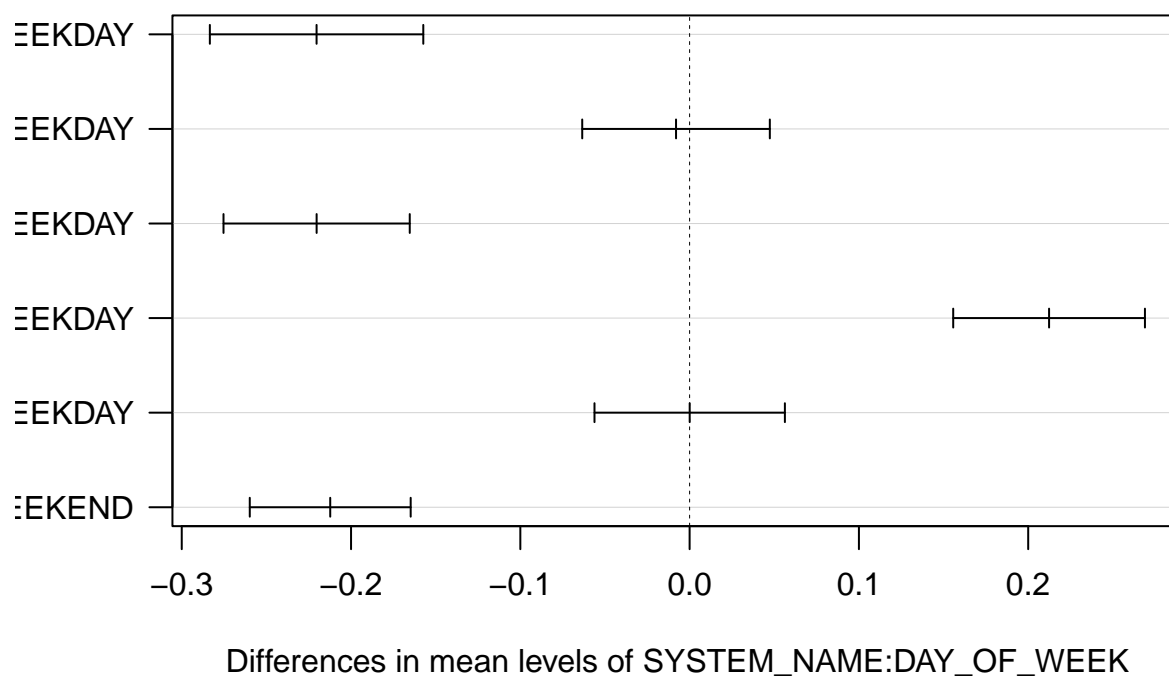
### 95% family-wise confidence level



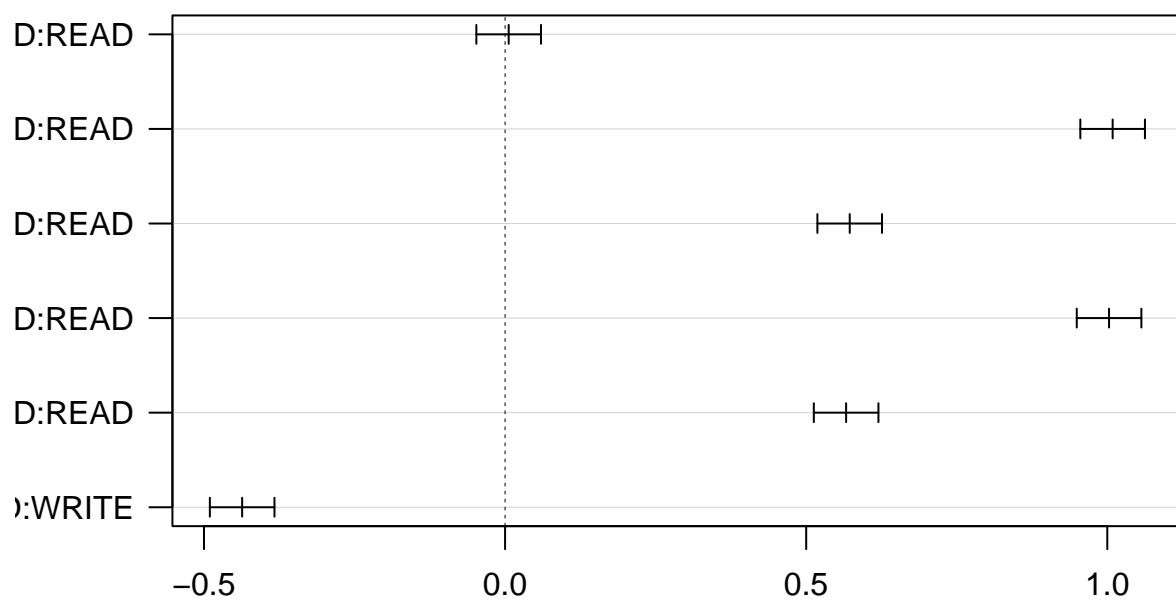
### 95% family-wise confidence level



### 95% family-wise confidence level



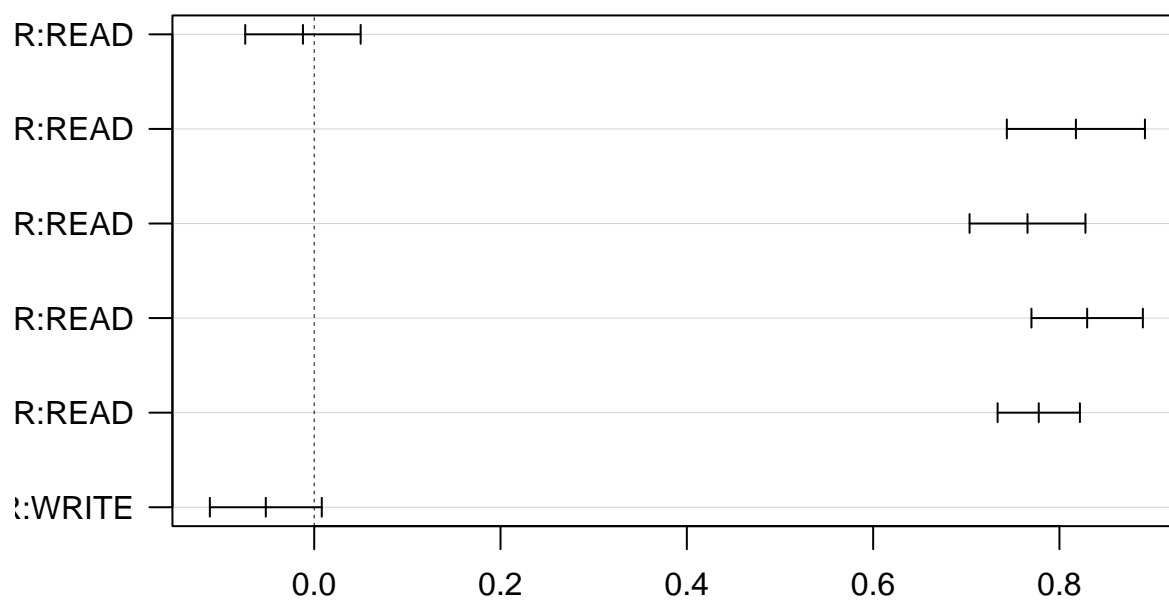
### 95% family-wise confidence level



Differences in mean levels of SYSTEM\_NAME:OPERATION\_TYPE

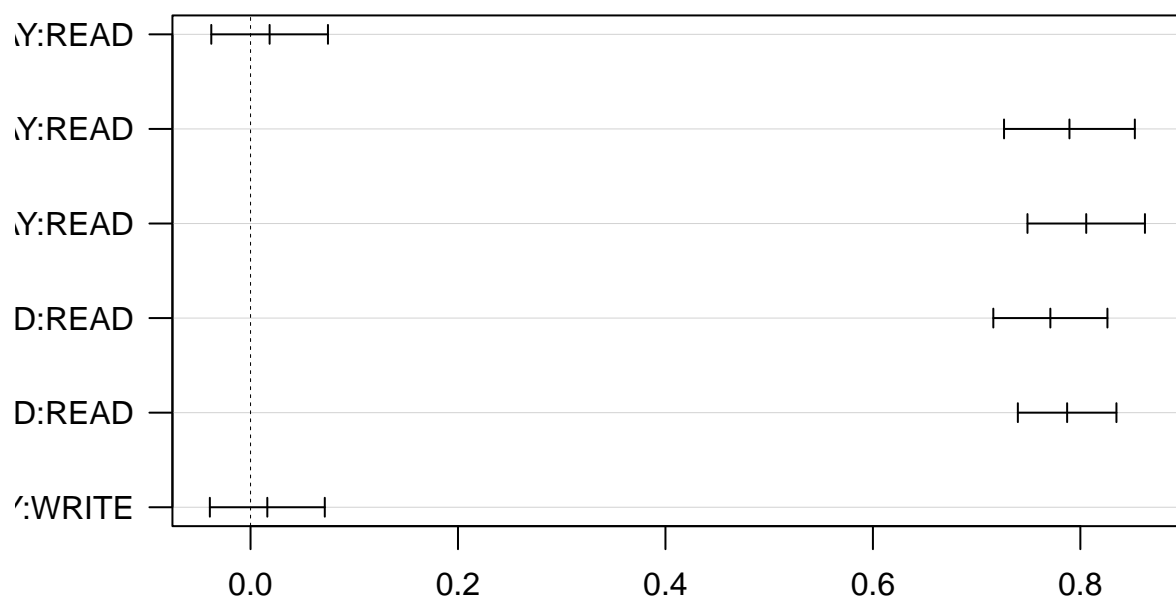


### 95% family-wise confidence level



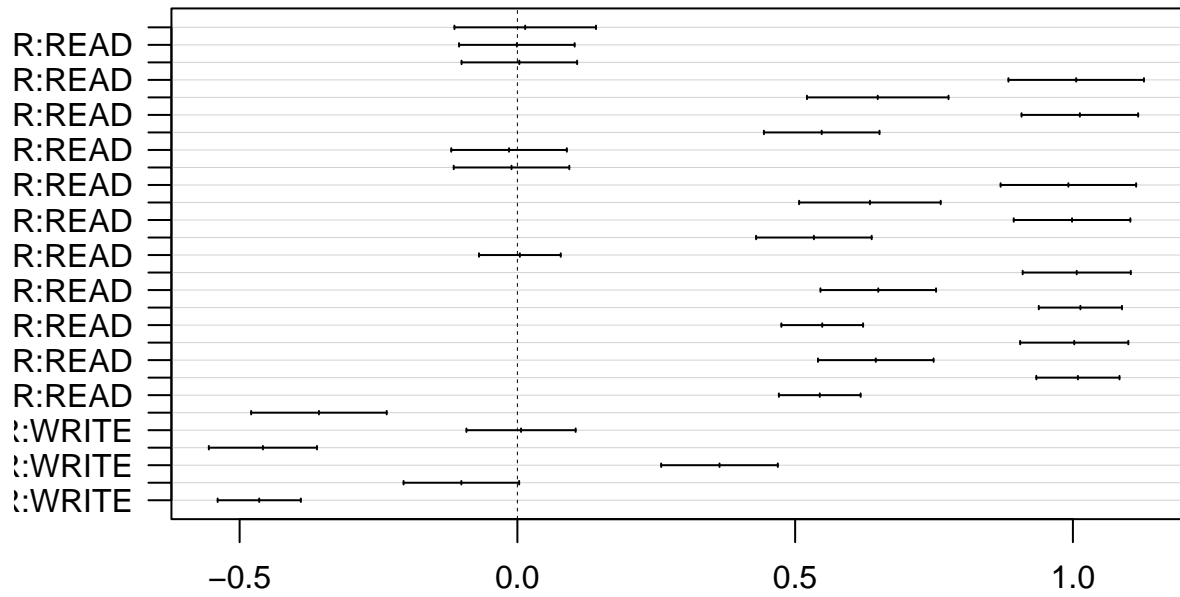
Differences in mean levels of TIME\_OF\_DAY:OPERATION\_TYPE

### 95% family-wise confidence level



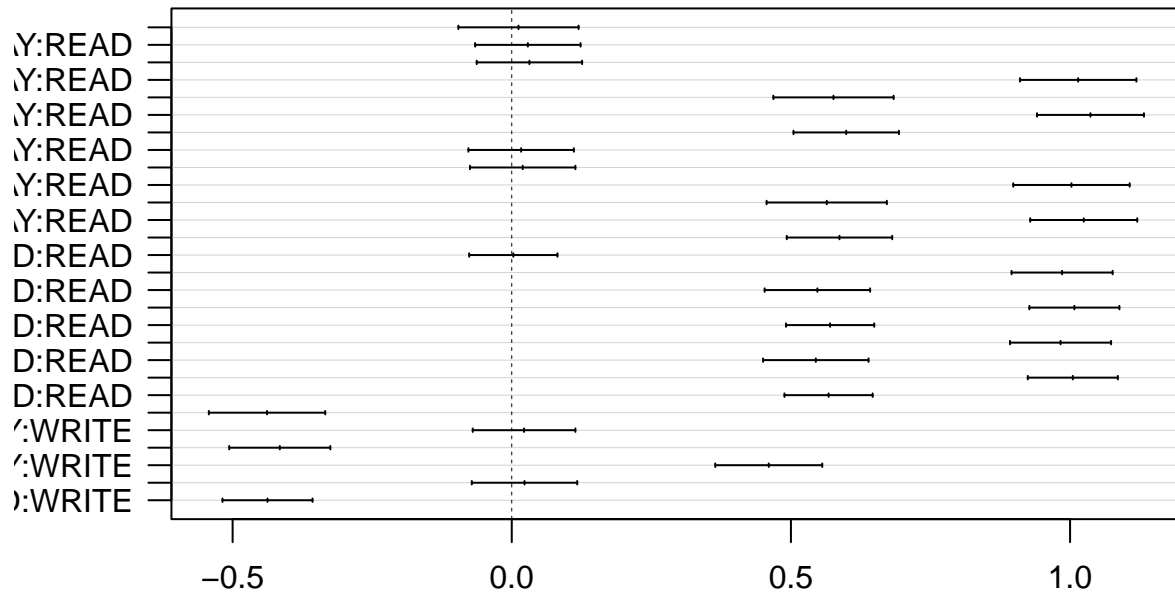
Differences in mean levels of DAY\_OF\_WEEK:OPERATION\_TYPE

### 95% family-wise confidence level



Differences in mean levels of SYSTEM\_NAME:TIME\_OF\_DAY:OPERATION\_TYPE

## 95% family-wise confidence level



Differences in mean levels of SYSTEM\_NAME:DAY\_OF\_WEEK: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. Day of the week 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.