
H2O.ai

Summary & Aggregation

```
 h2o.group_by(data, by, ..., gb.control =
list(na.methods = NULL, col.names = NULL))
```

Arguments

`data`	an H2OFrame object.
`by`	a list of column names
`gb.control`	a list of how to handle NA values in the dataset as well as how to name output columns
`. . .`	Any supported aggregated function: `mean, min, max, sum, sd, nrow`

Numeric Data Summaries

- `h2o_frame[x].cor(y=None, na_rm=False, use=None)`
- `h2o_frame[x].kurtosis(na_rm=False)`
- `h2o_frame[x].max()`
- `h2o_frame[x].mean(skipna=False)`
- `h2o_frame[x].median(na_rm=False)`
- `h2o_frame[x].min()`
- `h2o_frame[x].prod(na_rm=False)`
- `h2o_frame[x].quantile(…)`
- `h2o_frame[x].sd(na_rm=False)`
- `h2o_frame[x].skewness(na_rm=False)`
- `h2o_frame[x].sum(skipna=True)`
- `h2o_frame[x].var(y=None, na_rm=False, use=None)`

Summary & Aggregation

```
h2o.group_by(data, by, ..., gb.control =
list(na.methods = NULL, col.names = NULL))
```

Arguments

`data`	an H2OFrame object.
`by`	a list of column names
`gb.control`	a list of how to handle NA values in the dataset as well as how to name output columns
`. . .`	Any supported aggregated function: `mean, min, max, sum, sd, nrow`