DEEPHAVEN Data Labs

# Query Cheat Sheet

## Accessing and Creating Tables

### Getting Tables

```
db.t("Namespace", "Table")          //access historical data
db.i("Namespace", "Table")          //access intraday data
db.i("Namespace", "Table", isLive)  //access intraday data w/o ticking

db.getNamespaces()
db.getTableNames("Namespace")
db.getCatalog()    //get a table with available namespaces and table names
```

## Filtering Data

*Note: Backticks (`) are used for strings within strings.*

### where

```
t.where("Condition"…)
   //examples
   t.where("Date > `2018-01-01`")
   t.where("Timestamp <= '2018-01-01T16:00:00 NY'")
   t.where("USym = `AAPL`")
   t.where("USym in `AAPL`, `MSFT`, `GOOG`")        // filters a list of values
   t.where("USym icase in `aapl`, `msft`, `goog`")  //case insensitive
   t.where("USym not in `NVDA`, `AMZN`")

t1.whereIn(t2, "Match"…)          //filter to rows that match
t1.whereNotIn(t2, "Match"…)       //filter to rows that do not match
   //examples
   t1.whereIn(t2, "USym")
   t1.whereIn(t2, "USym=UnderlyingSym")    //can match different column names
```

### head/tail

```
t.head(rows)                    //first number of rows
t.headPct(decimalNum)           //first percentage of rows
t.headBy(rows, "Column"…)       //first number of rows for each subset
t.tail(rows) //last number of rows
t.tailPct(decimalNum)
t.tailBy(rows, "Column"…)
```

## Sorting Data

```
t.sort("Column"…)
t.sortDescending("Column"…)
t.reverse()                          //reverses order of entire table
t.restrictSortTo("Column"...)        //restricts sorting to specified columns
t.clearSortingRestrictions()         //removes sorting restrictions
```

## Data Selection

### In-memory

```
t.select("ColumnFormula"…)
t.update("ColumnFormula"…)        //evaluates data once and stores it in memory
```

### On-the-fly

```
t.view("ColumnFormula"…)
t.updateView("ColumnFormula"…)   //evaluates data but doesn't store results in memory
```

### Other Methods

```
t.selectDistinct("Column"…)      //only keep unique values
t.dropColumns("Column"…)
t.renameColumns("ColumnRename"…)
t.moveColumns(index, "Column"…)
t.moveUpColumns("Column"…)
```

## Grouping and Aggregating Data

### By and Ungroup

```
t.by("Column"…)
      //data in columns other than those specified are grouped into arrays

t.ungroup("Column"…)
      //distributes values in specified columns into arrays
      //each array value becomes its own row
```

### ByExternal

```
tables = t.byExternal("Column"…)
tables.getKeySet()           //returns array of keys
table = tables.get("Key")

//The following is required if the table broken down by multiple columns
import com.fishlib.datastructures.util.SmartKey
table = tables.get(new SmartKey("Key"…))
```

### Simple Aggregations

```
t.firstBy("Column"…)
t.lastBy("Column"…)
t.sumBy("Column"…)
t.avgBy("Column"…)
t.stdBy("Column"…)
t.varBy("Column"…)
t.medianBy("Column"…)
t.minBy("Column"…)
t.maxBy("Column"…)
t.countBy("CountColumn", "Column"…)
```

### Combined Aggregations

```
t.by(AggCombo(Agg…), "Column"…)

AggFirst("Column"…)
AggLast("Column"…)
AggSum("Column"…)
AggAvg("Column"…)
AggStd("Column"…)
AggVar("Column"…)
AggMedian("Column"…)
AggMin("Column"…)
AggMax("Column"…)
AggCount("Column")
```

```
//combined aggregation example
t.by(AggCombo(AggSum("Dollars", "Size"), AggAvg("AvgSize=Size")),"Symbol")
```

## Joining Data from Multiple Tables

```
t1.naturalJoin(t2, "Match…", "ColumnToAdd…")
t1.exactJoin(t2, "Match…", "ColumnToAdd…")
t1.join(t2, "Match…", "ColumnToAdd…")
t1.leftJoin(t2, "Match…", "ColumnToAdd…")

//as-of joins matching subsets need to be ordered
t1.aj(t2, "Match…, AjMatch", "ColumnToAdd…")
t1.raj(t2, "Match…, AjMatch", "ColumnToAdd…")
```

## Working with Time

*Note: Single quotes (') rather than backticks (`) are used within strings.*

### String Formats

```
yyyy-mm-ddThh:mm:ss.nanos TZ //date-time
#y#m#w#dT#h#m#s    //period
hh:mm:ss.nanos     //duration

//constants representing nanoseconds
threeSeconds = 3 * SECOND
fiveMinutes = 5 * MINUTE
twoHours = 2 * HOUR
sevenDays = 7 * DAY
```

### Useful Methods

```
currentDateNy()
lastBusinessDateNy()
formatDateNy(DateTime)    //converts to a date string
currentTime()
```

## Formatting Tables

### Numeric

```
t.formatColumns("ColumnName=Decimal(`.00#`)"…)
      //or Decimal(`\$.00#`) or Decimal(`0.00#`)
      //# means only display non-zero numbers in position
      //0 means display zeroes in position
```

### Color

```
t.formatColumns("ColumnName=Color")
t.formatColumnsWhere("ColumnName", "Condition", "Color")
t.formatRowWhere("Condition", "Color")
t.formatColumns("ColumnName=heatmap(Column, minValue, maxValue, minColor, maxColor)")
```

## Special Table Operations

### Table Tools

```
emptyTable(rows)
newTable(column…) //create table with columns specified below
merge(table…) //append table rows to one another with same columns
```

### Metadata

```
t.getMeta()
```