AGGREGATION

Aggregation

- `count():`
  - Number of documents in a collection
  - Number of results for a query

- `distinct():`
  - All distinct values for a key

  `db.runCommand("distinct" : "people", "key" : "age"})`

- `group():`
  - Group for each key value, with result document

Group

- Example: Stock Prices database

```
{"day" : "2010/10/03",
 "time" : "10/3/2010 03:57:01 GMT-400",
 "price" : 4.23}
{"day" : "2010/10/04",
 "time" : "10/4/2010 11:28:39 GMT-400",
 "price" : 4.27}
{"day" : "2010/10/03",
 "time" : "10/3/2010 05:00:23 GMT-400",
 "price" : 4.10}
{"day" : "2010/10/06",
 "time" : "10/6/2010 05:27:58 GMT-400",
 "price" : 4.30}
{"day" : "2010/10/04",
 "time" : "10/4/2010 08:34:50 GMT-400",
 "price" : 4.01}
```

116

116

Group

- Latest stock prices

```
[
  {"time" : "10/3/2010 05:00:23 GMT-400",
   "price" : 4.10},
  {"time" : "10/4/2010 11:28:39 GMT-400",
   "price" : 4.27},
  {"time" : "10/6/2010 05:27:58 GMT-400",
   "price" : 4.30}
]
```

117

117

Group

- Query

```
db.runCommand({"group" : {
  "ns" : "stocks",
  "key" : "day",
  "initial" : {"price" : null, "time" : 0},
  "$reduce" : function(doc, prev) {
    if (doc.time > prev.time) {
      prev.price = doc.price;
      prev.time = doc.time;
    }
  }
}})
```

Group

- Query

```
db.runCommand({"group" : {
  "ns" : "stocks",
  "key" : "day",
  "initial" : {"time" : 0},
  "$reduce" : function(doc, prev) {
    if (doc.time > prev.time) {
      prev.price = doc.price;
      prev.time = doc.time;
    }
  },
  "condition" : {"day" : {"$gt" : "2010/09/30"}}
}})
```

Group

- Query

```
db.runCommand({"group" : {
  "ns" : "stocks",
  "key" : "day",
  "initial" : {"time" : 0},
  "$reduce" : function(doc, prev) {
    if (doc.time > prev.time) {
      prev.price = doc.price;
      prev.time = doc.time;
    }
  },
  "condition" : {"day" : {"$gt" : "2010/09/30"}
                          {"$exists" : true}}
}})
```

120

120

Group

- Result

```
{
  "retval" :
    [ {
        "day" : "2010/10/04",
        "time" : "Mon Oct 04 2010 11:28:39 ..."
        "price" : 4.27
      },
      ...
    ],
  "count" : 734,
  "keys" : 30,
  "ok" : 1
}
```

121

121

4

Group

- Example: Blog posts

```
{
  "day" : "2010/10/12"
  "tags" : [ "sledding", "nosql" ]
}
{
  "day" : "2010/10/12"
  "tags" : [ "winter", "nosql" ]
}
{
  "day" : "2010/10/13"
  "tags" : [ "php" ]
}
...
```

122

Group

- Query: For blog site, find most popular tag for day

```
db.posts.group({
  "key" : "day",
  "initial" : {"tags" : {}},
  "$reduce" : function(doc, prev) {
      for (i in doc.tags) {
        if (doc.tags[i] in prev.tags)
{
          prev.tags[doc.tags[i]]++;
        } else {
          prev.tags[doc.tags[i]] = 1;
        }
      } ...
}})
```

prev.tags["nosql"]++
Is the same as
prev.tags.nosql++

123

Group

- Result
```
[
  {"day" : "2010/01/12",
   "tags" :
      {"nosql" : 4, "winter" : 10, "sledding" :
2}},
  {"day" : "2010/01/13",
   "tags" : {"soda" : 5, "php" : 2}},
  {"day" : "2010/01/14",
   "tags" :
      {"python" : 6, "winter" : 4, "nosql": 15}}
]
```

Group

- Query: For blog site, find most popular tag for day
```
db.posts.group({
  "key" : "day",
  "initial" : {"tags" : {}},
  "$reduce" : function(doc, prev) { ... }
  "finalize" : function(prev) {
    var mostPopular = 0;
    for (i in prev.tags) {
      if (prev.tags[i] > mostPopular) {
        prev.tag = i;
        mostPopular = prev.tags[i];
      }
    }
    delete prev.tags
}})
```

Group

- Result
  ```
  [
    {"day" : "2010/01/12", "tag" : "winter"},
    {"day" : "2010/01/13", "tag" : "soda"},
    {"day" : "2010/01/14", "tag" : "nosql"}
  ]
  ```

126

126

7