



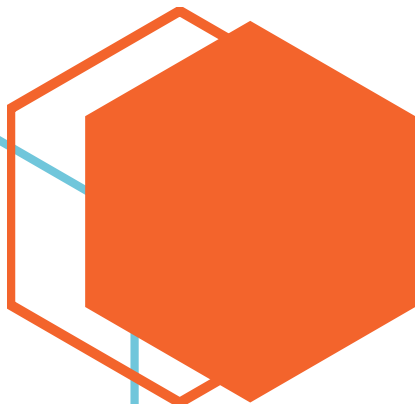
MongoDB Review CASESTUDY

19CSE357
BIGDATA ANALYTICS

CSE - E

DIVESH KOSURI
CB.EN.U4CSE19422

12-04-2022



FINDING THE MINIMUM AND MAXIMUM YEAR RECORDED IN THE DATABASE

QUERY 1: `db.delayData.find().sort({year:1}).limit(1).pretty();`

```
> db.delayData.find().sort({year:1}).limit(1).pretty();
{
  "_id" : ObjectId("625d88a82c8be4ce18263bf5"),
  "year" : 2003,
  "month" : 6,
  "carrier" : "AA",
  "carrier_name" : "American Airlines Inc.",
  "airport" : "ABQ",
  "airport_name" : "Albuquerque, NM: Albuquerque International Sunport",
  "arr_flights" : 307,
  "arr_del15" : 56,
  "carrier_ct" : 14.68,
  "weather_ct" : 10.79,
  "nas_ct" : 19.09,
  "security_ct" : 1.48,
  "late_aircraft_ct" : 9.96,
  "arr_cancelled" : 1,
  "arr_diverted" : 1,
  "arr_delay" : 2530,
  "carrier_delay" : 510,
  "weather_delay" : 621,
  "nas_delay" : 676,
  "security_delay" : 25,
  "late_aircraft_delay" : 698,
  "" : ""
}
```

QUERY 2: `db.delayData.find().sort({year:-1}).limit(1).pretty();`

```
> db.delayData.find().sort({year:-1}).limit(1).pretty();
{
  "_id" : ObjectId("625d88cb2c8be4ce1828be69"),
  "year" : 2013,
  "month" : 1,
  "carrier" : "9E",
  "carrier_name" : "Pinnacle Airlines Inc.",
  "airport" : "DFW",
  "airport_name" : "Dallas/Fort Worth, TX: Dallas/Fort Worth International",
  "arr_flights" : 139,
  "arr_del15" : 37,
  "carrier_ct" : 13.04,
  "weather_ct" : 0.8,
  "nas_ct" : 13.92,
  "security_ct" : 0,
  "late_aircraft_ct" : 9.24,
  "arr_cancelled" : 2,
  "arr_diverted" : 5,
  "arr_delay" : 1860,
  "carrier_delay" : 789,
  "weather_delay" : 75,
  "nas_delay" : 451,
  "security_delay" : 0,
  "late_aircraft_delay" : 545,
  "" : ""
}
```

FINDING THE YEARS OF SERVICE [AGGREGATE FUNTIONS]

QUERY 3: `db.delayData.aggregate([{$group : {_id : "$carrier", MinYear : {$min : "$year"}}}]);`

```
> db.delayData.aggregate([{$group : {_id : "$carrier", MinYear : {$min : "$year"}}}]);
{ "_id" : "VX", "MinYear" : 2012 }
{ "_id" : "XE", "MinYear" : 2006 }
{ "_id" : "AQ", "MinYear" : 2006 }
{ "_id" : "F9", "MinYear" : 2005 }
{ "_id" : "OH", "MinYear" : 2004 }
{ "_id" : "HA", "MinYear" : 2003 }
{ "_id" : "9E", "MinYear" : 2007 }
{ "_id" : "WN", "MinYear" : 2003 }
{ "_id" : "US", "MinYear" : 2003 }
{ "_id" : "AA", "MinYear" : 2003 }
{ "_id" : "AS", "MinYear" : 2003 }
{ "_id" : "TZ", "MinYear" : 2003 }
{ "_id" : "DH", "MinYear" : 2003 }
{ "_id" : "B6", "MinYear" : 2003 }
{ "_id" : "UA", "MinYear" : 2003 }
{ "_id" : "CO", "MinYear" : 2003 }
{ "_id" : "MQ", "MinYear" : 2003 }
{ "_id" : "EV", "MinYear" : 2003 }
{ "_id" : "DL", "MinYear" : 2003 }
{ "_id" : "NW", "MinYear" : 2003 }
Type "it" for more
```

QUERY 4: `db.delayData.aggregate([{$group : {_id : "$carrier", MaxYear : {$max : "$year"}}}]);`

```
> db.delayData.aggregate([{$group : {_id : "$carrier", MaxYear : {$max : "$year"}}}]);
{ "_id" : "VX", "MaxYear" : 2013 }
{ "_id" : "XE", "MaxYear" : 2011 }
{ "_id" : "AQ", "MaxYear" : 2008 }
{ "_id" : "F9", "MaxYear" : 2013 }
{ "_id" : "OH", "MaxYear" : 2010 }
{ "_id" : "HA", "MaxYear" : 2013 }
{ "_id" : "9E", "MaxYear" : 2013 }
{ "_id" : "WN", "MaxYear" : 2013 }
{ "_id" : "US", "MaxYear" : 2013 }
{ "_id" : "AA", "MaxYear" : 2013 }
{ "_id" : "AS", "MaxYear" : 2013 }
{ "_id" : "TZ", "MaxYear" : 2006 }
{ "_id" : "DH", "MaxYear" : 2005 }
{ "_id" : "B6", "MaxYear" : 2013 }
{ "_id" : "UA", "MaxYear" : 2013 }
{ "_id" : "CO", "MaxYear" : 2011 }
{ "_id" : "MQ", "MaxYear" : 2013 }
{ "_id" : "EV", "MaxYear" : 2013 }
{ "_id" : "DL", "MaxYear" : 2013 }
{ "_id" : "NW", "MaxYear" : 2009 }
Type "it" for more
```

TO FIND THE TOTAL DELAY OCCURRED DUE TO CARRIER LOAD-UNLOAD

Query 5: `db.delayData.aggregate([{$group : {_id : "$carrier", TotalDelay: {$sum : "$carrier_ct"}}}]);`

```
> db.delayData.aggregate([{$group : {_id : "$carrier", TotalDelay: {$sum : "$carrier_ct"}}}]);
{ "_id" : "VX", "TotalDelay" : 2740.08 }
{ "_id" : "XE", "TotalDelay" : 102311.54 }
{ "_id" : "AQ", "TotalDelay" : 3257.98 }
{ "_id" : "F9", "TotalDelay" : 37074.63 }
{ "_id" : "OH", "TotalDelay" : 146193.59 }
{ "_id" : "HA", "TotalDelay" : 27486.12 }
{ "_id" : "9E", "TotalDelay" : 65592.78 }
{ "_id" : "WN", "TotalDelay" : 523048.34 }
{ "_id" : "US", "TotalDelay" : 229268.93 }
{ "_id" : "AA", "TotalDelay" : 367637.12 }
{ "_id" : "AS", "TotalDelay" : 88999.15 }
{ "_id" : "TZ", "TotalDelay" : 6563.67 }
{ "_id" : "DH", "TotalDelay" : 26079.83 }
{ "_id" : "B6", "TotalDelay" : 103185.42 }
{ "_id" : "UA", "TotalDelay" : 243557.23 }
{ "_id" : "CO", "TotalDelay" : 112686.43 }
{ "_id" : "MQ", "TotalDelay" : 266857.85 }
{ "_id" : "EV", "TotalDelay" : 294554.02 }
{ "_id" : "DL", "TotalDelay" : 312964.66 }
{ "_id" : "NW", "TotalDelay" : 198141.92 }
```

INSERTING NEW RECORD

QUERY 6: `db.delayDB.insert([{"year" : 2022, "month" : 4, "carrier" : "AAA", "carrier_name" : "Divesh Airlines Inc.", "airport" : "DSM", "airport_name" : "Divesh International", "arr_flights" : 422, "arr_del15" : 15, "carrier_ct" : 3.39, "weather_ct" : 1.57, "nas_ct" : 5.07, "security_ct" : 0, "late_aircraft_ct" : 6.97, "arr_cancelled" : 0, "arr_diverted" : 0, "arr_delay" : 0, "carrier_delay" : 212, "weather_delay" : 50, "nas_delay" : 259, "security_delay" : 0, "late_aircraft_delay" : 345, "" : ""}]);`

```
> db.delayDB.insert([{"year" : 2022, "month" : 4, "carrier" : "AAA", "carrier_name" : "Divesh Airlines Inc.", "airport" : "DSM", "airport_name" : "Divesh International", "arr_flights" : 422, "arr_del15" : 15, "carrier_ct" : 3.39, "weather_ct" : 1.57, "nas_ct" : 5.07, "security_ct" : 0, "late_aircraft_ct" : 6.97, "arr_cancelled" : 0, "arr_diverted" : 0, "arr_delay" : 0, "carrier_delay" : 212, "weather_delay" : 50, "nas_delay" : 259, "security_delay" : 0, "late_aircraft_delay" : 345, "" : ""}]);
BulkWriteResult({
  "writeErrors" : [ ],
  "writeConcernErrors" : [ ],
  "nInserted" : 1,
  "nUpserted" : 0,
  "nMatched" : 0,
  "nModified" : 0,
  "nRemoved" : 0,
  "upserted" : [ ]
})
```

MAP-REDUCE QUERY 1

QUERY 7: Defining Mapper and Reducer

- `var mapCheck1 = function(){ emit(this.carrier_name, this.carrier_delay);}`
- `var reduceCheck1 = function(key, valArr) {return Array.sum(valArr);}`
- `db.delayData.mapReduce(mapCheck1, reduceCheck1,{out:"outputMR"});`

```
> var mapCheck1 = function(){ emit( this.carrier_name, this.carrier_delay );}
> var reduceCheck1 = function( key, valArr ) {return Array.sum( valArr );}
> db.delayData.mapReduce( mapCheck1, reduceCheck1,{out:"outputMR"});
{
  "result" : "outputMR",
  "timeMillis" : 2446,
  "counts" : {
    "input" : 173657,
    "emit" : 173657,
    "reduce" : 4272,
    "output" : 25
  },
  "ok" : 1
}
```

QUERY 8: Loading the output with sorted order according to the requirement:

```
> db.outputMR.find().sort({year:-1});
{ "_id" : "ATA Airlines d/b/a ATA", "value" : 405504 }
{ "_id" : "AirTran Airways Corporation", "value" : 4028057 }
{ "_id" : "Alaska Airlines Inc.", "value" : 5062626 }
{ "_id" : "Aloha Airlines Inc.", "value" : 144922 }
{ "_id" : "America West Airlines Inc.", "value" : 1699448 }
{ "_id" : "American Airlines Inc.", "value" : 22870499 }
{ "_id" : "American Eagle Airlines Inc.", "value" : 16095940 }
{ "_id" : "Atlantic Coast Airlines", "value" : 1064001 }
{ "_id" : "Atlantic Southeast Airlines", "value" : 13837566 }
{ "_id" : "Comair Inc.", "value" : 8415020 }
{ "_id" : "Continental Air Lines Inc.", "value" : 6480851 }
{ "_id" : "Delta Air Lines Inc.", "value" : 19040800 }
{ "_id" : "ExpressJet Airlines Inc.", "value" : 14492140 }
{ "_id" : "Frontier Airlines Inc.", "value" : 1721862 }
{ "_id" : "Hawaiian Airlines Inc.", "value" : 1355530 }
{ "_id" : "Independence Air", "value" : 304026 }
{ "_id" : "JetBlue Airways", "value" : 6512832 }
{ "_id" : "Mesa Airlines Inc.", "value" : 8681157 }
{ "_id" : "Northwest Airlines Inc.", "value" : 11745027 }
{ "_id" : "Pinnacle Airlines Inc.", "value" : 4365161 }
Type "it" for more
> it
{ "_id" : "SkyWest Airlines Inc.", "value" : 17016279 }
{ "_id" : "Southwest Airlines Co.", "value" : 22184643 }
{ "_id" : "US Airways Inc.", "value" : 11378650 }
{ "_id" : "United Air Lines Inc.", "value" : 14375218 }
{ "_id" : "Virgin America", "value" : 154025 }
```

MAP-REDUCE QUERY 2

QUERY 9: Defining Mapper and Reducer

```
var mapCheck2 = function(){ emit( this.carrier_name, this.weather_delay );}
var reduceCheck2 = function( key, valArr ) {return Array.sum( valArr );}
db.delayData.mapReduce( mapCheck2, reduceCheck2, {out:"NewOutputMR"});
```

```
> var mapCheck2 = function(){ emit( this.carrier_name, this.weather_delay );}
> var reduceCheck2 = function( key, valArr ) {return Array.sum( valArr );}
> db.delayData.mapReduce( mapCheck2, reduceCheck2, {out:"NewOutputMR"});
{
  "result" : "NewOutputMR",
  "timeMillis" : 992,
  "counts" : {
    "input" : 173657,
    "emit" : 173657,
    "reduce" : 4272,
    "output" : 25
  },
  "ok" : 1
}
```

QUERY 10: Loading the output with sorted order according to the requirement:

```
> db.NewOutputMR.find().sort({year:-1})
{ "_id" : "ATA Airlines d/b/a ATA", "value" : 23136 }
{ "_id" : "AirTran Airways Corporation", "value" : 306101 }
{ "_id" : "Alaska Airlines Inc.", "value" : 345748 }
{ "_id" : "Aloha Airlines Inc.", "value" : 2297 }
{ "_id" : "America West Airlines Inc.", "value" : 71713 }
{ "_id" : "American Airlines Inc.", "value" : 5228234 }
{ "_id" : "American Eagle Airlines Inc.", "value" : 4157053 }
{ "_id" : "Atlantic Coast Airlines", "value" : 317277 }
{ "_id" : "Atlantic Southeast Airlines", "value" : 4318712 }
{ "_id" : "Comair Inc.", "value" : 4849900 }
{ "_id" : "Continental Air Lines Inc.", "value" : 1414448 }
{ "_id" : "Delta Air Lines Inc.", "value" : 2215918 }
{ "_id" : "ExpressJet Airlines Inc.", "value" : 2546530 }
{ "_id" : "Frontier Airlines Inc.", "value" : 183900 }
{ "_id" : "Hawaiian Airlines Inc.", "value" : 23615 }
{ "_id" : "Independence Air", "value" : 34186 }
{ "_id" : "JetBlue Airways", "value" : 635836 }
{ "_id" : "Mesa Airlines Inc.", "value" : 952900 }
{ "_id" : "Northwest Airlines Inc.", "value" : 2342139 }
{ "_id" : "Pinnacle Airlines Inc.", "value" : 669742 }
```