1. **Heatmap Across the US (Easy)**

**// Description: selecting the latitude and longitude information**

db['f3'].aggregate([{

"$project" : {

"latitude" : "$location\_info.latitude",

"longitude" : "$location\_info.longitude",

"\_id":0

}}])

1. **Scatterplot # Guns involved vs. Casualties (Easy)**

**// Description: comparing # guns involved and casualties**

db.f3.aggregate([

{"$match":

{'$and': [

{'gun\_info.n\_guns\_involved': {'$ne':null}},

{'gun\_info.n\_guns\_involved': {'$lt' : 20}}]

}},

{"$project": {

"totalCasualties" : {"$add" : ["$casualties.n\_killed" , "$casualties.n\_injured"]},

"n\_guns\_involved" : "$gun\_info.n\_guns\_involved",

"\_id":0

}}

])

1. **Number of incidents for 2014-2017 (Easy)**

**// Description: Showing trend for # of incidents across the US for 2014-2017**

db.f3.aggregate([

{$project: {

year: {$year : '$date' }

}},

{$bucket: {

groupBy: "$year",

boundaries: [2014, 2015, 2016, 2017, 2018],

default:0

}},

{$match: {'\_id': {'$ne':0}}}

])

1. **Incidents within 20 km of Atlanta (Easy)**

**// Description: clustering incidents within 20km of Atlanta using searched up Atlanta coordinates**

db.f3.aggregate([

{$geoNear: {

near: {type: "Point", coordinates: [-84.386330, 33.753746]},

maxDistance: 20000,

distanceField: 'location\_info.geo',

spherical : true

}},

{$project: {

latitude : "$location\_info.latitude",

longitude: "$location\_info.longitude",

date : 1,

\_id : 0

}},

])

1. **Months with the Most Gun Violence (Easy)**

**// Description: This query identifies which months of the year tend to have the most gun violence**

db['f3'].aggregate([

{

$project: {

month: {$dateToString : {format : '%m', date: '$date'}}

}

},

{

$group: {

'\_id': '$month',

count: {$sum:1}

}

},

{

$sort: {

count: -1

}

}

])

1. **State Dangerous Ranking (Medium)**

**// Description: Rank how dangerous each state is based on the total killed and then total injured**

db.f3.aggregate([

{"$group" : {

"\_id" : {"state": "$location\_info.state"},

"total\_injured": {$sum: "$casualties.n\_injured"},

"total\_killed": {$sum:"$casualties.n\_killed"}

}},

{"$match":{

$and:[{"total\_killed":{$gt:500}},

{"total\_injured":{$gt:700}}]}},

{"$project" : {

"state" : "$\_id.state",

"total\_injured":1,

"total\_killed":1,

"\_id" : 0

}},

{$sort: {total\_killed:-1, total\_injured:-1}}

])

1. **Chicago # incidents in 2017 (Medium)**

**// Description: Show the trend for total casualties and # incidents in 2017 in Chicago**

db.f3.aggregate([

{"$project": {

"year": { "$year": "$date" },

"location\_info":1,

"date":1,

"casualties":1

}},

{"$match" : {

"$and":[{

"year" : {'$eq': 2017},

"location\_info.city\_or\_county":"Chicago"

}]

}},

{

"$project": {

"month": {"$dateToString" : {format : '%m', date: '$date'}},

"location\_info":1,

"date":1,

"casualties":1

}},

{"$group":{

"\_id":{"month":"$month"},

"total\_killed":{$sum:"$casualties.n\_killed"},

"total\_injured":{$sum:"$casualties.n\_injured"},

"num\_incidents": {$push: '$\_id'}

}},

{"$project":{

"month":"$\_id.month",

"total\_casualties":{$sum:["$total\_killed", "total\_injured"]},

"num\_incidents":{"$size":"$num\_incidents"}

"\_id":0

}},

{"$sort":{"month":1}}

])

1. **Gender Ratio of Deaths (Medium)**

**// Description: the percentages of males and females among the killed**

db.f3.aggregate([

{

$project: {

'participant\_info.participant\_gender' : 1,

'participant\_info.participant\_status' : 1,

'\_id' : 0

}

},

{

$unwind: '$participant\_info'

},

{

$match: {

'$or' : [

{

'$and': [

{'participant\_info.participant\_gender' : 'Male'},

{'participant\_info.participant\_status' : 'Killed'}

]},

{

'$and': [

{'participant\_info.participant\_gender' : 'Female'},

{'participant\_info.participant\_status' : 'Killed'}

]}

]

}

},

{

$group: {

\_id: '$participant\_info.participant\_gender',

total : {$sum: 1},

}

},

{

$project: {

'\_id': 1,

'total\_killings' : '$total',

}

},

])

1. **Average Age of Cities with Most Incidents (Medium)**

**// Description: the average age those involved of the top 5 cities with the most incidents**

db.f3.aggregate([

{

$project: {

'participant\_info.participant\_age' : 1,

'location\_info.city\_or\_county' : 1,

'\_id' : 0

}

},

{

$unwind: '$participant\_info'

},

{

$match : {

'participant\_info.participant\_age': {'$ne':null},

}

},

{

$group: {

\_id: '$location\_info.city\_or\_county',

count: {$sum:1},

ages: {$push: '$participant\_info.participant\_age'},

average : {$avg: '$participant\_info.participant\_age'}

}

},

{

$sort: {count : -1}

},

{

$limit : 5

},

{

$project: {

\_id : 1,

average : 1,

}

}

])

1. **Average # of people injured/killed vs. officer involvement (Medium)**

**// Description: This query is comparing the average number of people killed and injured when there is an officer present and when there is not an officer present**

db.f3.aggregate([

{

$project:{

"\_id":1,

"incident":"$misc.incident\_characteristics",

"nInjured": "$casualties.n\_injured",

"nKilled": "$casualties.n\_killed"

}

},

{

$unwind:"$incident"

},

{

$project:{

"\_id":1,

"officer":{ $regexMatch:{

input: "$incident", regex:/.\*Officer.\*/, options:"i"

}},

"nInjured": "$nInjured",

"nKilled": "$nKilled",

}

},

{

$group: {

"\_id": "$officer",

avgInjured: { $avg: "$nInjured" },

avgKilled: { $avg: "$nKilled" }

}

}

])

1. **Cities with the Highest Rate of Gang-related Gun Incidents (Medium)**

**// Description: This query identifies which cities have the most gun violence that is related to gang activity.**

db['f3'].aggregate([

{

$project: {

'location\_info.city\_or\_county':1,

'misc.incident\_characteristics':1,

'\_id':0

}

},

{

$unwind: '$misc.incident\_characteristics'

},

{

$match: {

'misc.incident\_characteristics': {

$all: ['Gang involvement']

}

}

},

{

$group: {

'\_id': '$location\_info.city\_or\_county',

count: {$sum:1}

}

},

{

$sort: {

count: -1

}

}

])

1. **Gun violence by day of the year in the city with the most gun (Chicago) violence incidents (Medium)**

**// Description: This query analyzes the number of incidents per day of the year in the city with the most gun violence. The goal was to analyze trends throughout the year in which the number of incidents may spike or slow down.**

db['f3'].aggregate([

{

$match: { 'location\_info.city\_or\_county': 'Chicago' }

},

{

$project: {

dayOfYear: {$dateToString : {format : '%j', date: '$date'}}

}

},

{

$group: {

'\_id': '$dayOfYear',

count: {$sum:1}

}

},

{

$sort: {

\_id:1

}

}

])

1. **Domestic violence Ranking (Hard)**

**// Description: Ranking of the cities in terms of number of domestic violence**

db.f3.aggregate([

{$project:{

"\_id":1,

"incident":"$misc.incident\_characteristics",

"city\_or\_county":"$location\_info.city\_or\_county"

}},

{$unwind:"$incident"},

{$project:{

"\_id":1,

"city\_or\_county":1,

"domestic":{$regexMatch:{input: "$incident", regex:/.\*domestic.\*/, options:"i"}}

}},

{$match:{"domestic":true}},

{$group:{

"\_id":{ "city\_or\_county":"$city\_or\_county"},

"incident\_id":{$addToSet:"$\_id"}

}},

{$project:{

"city\_or\_county":"$\_id.city\_or\_county",

"incident\_num":{$size:"$incident\_id"},

"\_id":0

}},

{$sort:{incident\_num:-1}},

{$limit:10}

])

1. **Ranking deaths and Injuries by congressional districts in Georgia (Hard)**

**// Description: There are 14 congressional districts in GA (for the 14 US House of Representatives), and we are ranking them to see which districts are the most deadly and which districts are the least deadly.**

db.f3.aggregate([

{ $match: { $and: [ { "location\_info.state": "Georgia" }, { 'law\_info.congressional\_district': {'$ne':null} } ] } },

{

$group : {

"\_id" : {

"districtNum" : "$law\_info.congressional\_district",

},

"total\_killed":{$sum:"$casualties.n\_killed"},

"total\_injured": {$sum: "$casualties.n\_injured"}

}

},

{

$project : {

"total\_killed":1,

"total\_injured":1,

"\_id" : 0,

"district" :

{

$switch:

{

branches: [

{

case: { $eq: [ "$\_id.districtNum", 1 ] },

then: "Pooler"

},

{

case: { $eq: [ "$\_id.districtNum", 2 ] },

then: "Albany"

},

{

case: { $eq: [ "$\_id.districtNum", 3 ] },

then: "West Point"

},

{

case: { $eq: [ "$\_id.districtNum", 4 ] },

then: "Lithonia"

},

{

case: { $eq: [ "$\_id.districtNum", 5 ] },

then: "North Atlanta"

},

{

case: { $eq: [ "$\_id.districtNum", 6 ] },

then: "Marietta"

},

{

case: { $eq: [ "$\_id.districtNum", 7 ] },

then: "Peachtree Corners"

},

{

case: { $eq: [ "$\_id.districtNum", 8 ] },

then: "Tifton"

},

{

case: { $eq: [ "$\_id.districtNum", 9 ] },

then: "Gainesville"

},

{

case: { $eq: [ "$\_id.districtNum", 10 ] },

then: "Bethlehem"

},

{

case: { $eq: [ "$\_id.districtNum", 11 ] },

then: "Cassville"

},

{

case: { $eq: [ "$\_id.districtNum", 12 ] },

then: "Augusta"

},

{

case: { $eq: [ "$\_id.districtNum", 13 ] },

then: "South Atlanta"

},

{

case: { $eq: [ "$\_id.districtNum", 14 ] },

then: "Ranger"

}

],

default: "No congressional district found."

}

}

}

},

{

$sort: {total\_killed:-1}

}

]);

1. **Female Suspects Across State (Hard)**

**// Description: compare the number of incidents that had female suspects across the states**

db.f3.aggregate([

{$project:{"participant\_info.participant\_gender":1,

"participant\_info.participant\_type":1,

"state":"$location\_info.state"}},

{$unwind:"$participant\_info"},

{$match:{"$and" : [ {

"participant\_info.participant\_gender":"Female"

}, {

"participant\_info.participant\_type" : "Subject-Suspect"

}]}},

{$group:{

"\_id":{ "state":"$state"},

"incident":{$addToSet:"$\_id"},

"state":{$addToSet:"$state"}

}},

{$project:{

"state":1,

"incident\_num":{$size:"$incident"},

"\_id":0

}},

{$unwind:"$state"},

{$sort:{incident\_num:-1}},

{ "$limit" : 10 }

])

1. **Who’s usually injured & killed in Home Invasion (Hard)**

**// Description: This query is aimed at identifying who is most often injured and killed in home invasions. Is it the suspect or the resident? This query has been written twice, once using the MapReduce functionality and the other time using aggregation.**

**MapReduce**

var mapFunc = function() {

for (i = 0; i < this.misc.incident\_characteristics.length; i++) {

var characteristic = this.misc.incident\_characteristics[i]

if (characteristic.includes("Home Invasion ")) {

emit(characteristic, 1)

}

}

}

var reduceFunc = function(characteristic, values) {

count = 0

for (i = 0; i < values.length; i++) {

count += values[i]

}

return count

}

db['f3'].mapReduce(mapFunc, reduceFunc, {out: 'characteristic\_stats'})

db.characteristic\_stats.find()

**Mongo/Knowi**

db['f3'].aggregate([

{$project: {'Characteristics': '$misc.incident\_characteristics'}},

{$unwind: '$Characteristics'},

{$facet: {'Home Invasion - No death or injury': [

{$project: {

\_id: 1,

'homeInvasion': { $regexMatch:

{ input: '$Characteristics',

regex:/.\*Home Invasion - No death or injury.\*/,

options:"i"}},

'characteristic': '$Characteristics'

}},

{$match: {homeInvasion: true}},

{$count: 'count'},

{$set:{"Name": "No death of injury"}}

],

'Home Invasion - Resident injured': [

{$project: {

\_id: 1,

'homeInvasion': { $regexMatch: { input: '$Characteristics',

regex:/.\*Home Invasion - Resident injured.\*/,

options:"i"}},

'characteristic': '$Characteristics'

}},

{$match: {homeInvasion: true}},

{$count: 'count'},

{$set:{"Name": "Resident injured"}}

],

'Home Invasion - Resident killed': [

{$project: {

\_id: 1,

'homeInvasion': { $regexMatch: { input: '$Characteristics',

regex:/.\*Home Invasion - Resident killed.\*/,

options:"i"}},

'characteristic': '$Characteristics'

}},

{$match: {homeInvasion: true}},

{$count: 'count'},

{$set:{"Name": "Resident killed"}}

],

'Home Invasion - subject/suspect/perpetrator injured': [

{$project: {

\_id: 1,

'homeInvasion': { $regexMatch: { input: '$Characteristics',

regex:/.\*Home Invasion - subject\/suspect\/perpetrator injured.\*/,

options:"i"}},

'characteristic': '$Characteristics'

}},

{$match: {homeInvasion: true}},

{$count: 'count'},

{$set:{"Name": "subject/suspect/perpetrator injured"}}

],

'Home Invasion - subject/suspect/perpetrator killed': [

{$project: {

\_id: 1,

'homeInvasion': { $regexMatch: { input: '$Characteristics',

regex:/.\*Home Invasion - subject\/suspect\/perpetrator killed.\*/,

options:"i"}},

'characteristic': '$Characteristics'

}},

{$match: {homeInvasion: true}},

{$count: 'count'},

{$set:{"Name": "subject/suspect/perpetrator killed"}}

]

}

},

{'$project': { 'Home Invasion': {

"$setUnion": ['$Home Invasion - No death or injury',

'$Home Invasion - Resident injured',

'$Home Invasion - Resident killed',

'$Home Invasion - subject/suspect/perpetrator injured',

'$Home Invasion - subject/suspect/perpetrator killed'] }}},

{"$unwind":"$Home Invasion"},

{"$project":{"Count":"$Home Invasion.count",

"Type":"$Home Invasion.Name"}}

])

1. **Box Plot of Top 5 most incident cities**

**// Description: gives the 5 number summary (min, max, median, 1st and 3rd quartile) of the Top 5 cities with the most incidents, and results in a box plot**

db.f3.aggregate([

{

$project: {

'participant\_info.participant\_age' : 1,

'location\_info.city\_or\_county' : 1,

'\_id' : 0

}

},

{

$unwind: '$participant\_info'

},

{

$match : {

$and : [{

'participant\_info.participant\_age': {'$ne':null},

'participant\_info.participant\_age': {'$lt':100},

}]

}

},

{

$group: {

\_id: '$location\_info.city\_or\_county',

count: {$sum:1},

ages: {$push: '$participant\_info.participant\_age'},

}

},

{

$sort: {count : -1}

},

{

$limit : 5

},

{

$project: {

\_id : 1,

ages: 1,

size : {'$size' : '$ages'},

max: {$max : '$ages'},

min: {$min : '$ages'}

}

},

{

$unwind: '$ages'

},

{

$sort: {ages: 1}

},

{

$group: {

\_id: '$\_id',

ages: {$push: '$ages'},

}

},

{

$project: {

ages: 1,

min: {$min: '$ages'},

max: {$max: '$ages'},

isEven: {$eq: [{$mod: [{$size : '$ages'}, 2]}, 0]},

mid: {$trunc : { '$divide' : [{$size : '$ages'}, 2]}},

first\_q: {$trunc : {$divide : [{$size : '$ages'}, 4]}},

third\_q: {$trunc : {$multiply : [3, {$divide : [{$size : '$ages'}, 4]}]}},

}

},

{

$project: {

\_id: 1,

ages: 1,

min: 1,

max: 1,

isEven: 1,

mid: 1,

first\_q:1,

third\_q: 1,

even\_mid: {$subtract: ['$mid', 1]},

even\_first: {$subtract: ['$first\_q', 1]},

even\_third: {$subtract: ['$third\_q', 1]},

mid\_val: {$arrayElemAt: ['$ages', '$mid']},

first\_val: {$arrayElemAt: ['$ages', '$first\_q']},

third\_val: {$arrayElemAt: ['$ages', '$third\_q']},

}

},

{

$project: {

\_id: 1,

ages: 1,

min: 1,

max: 1,

isEven: 1,

mid: 1,

first\_q:1,

third\_q: 1,

mid\_val: 1,

first\_val: 1,

third\_val: 1,

even\_mid: {$arrayElemAt: ['$ages', '$even\_mid']},

even\_first: {$arrayElemAt: ['$ages', '$even\_first']},

even\_third: {$arrayElemAt: ['$ages', '$even\_third']},

}

},

{

$project: {

\_id: 1,

min: 1,

max: 1,

median: {

$cond: {

if : '$isEven',

then: {$avg: ['$mid\_val', '$even\_mid']},

else: {$add : ['$mid\_val', 0]}

}

},

first\_q: {

$avg : ['$first\_val', '$even\_first']

},

third\_q: {

$avg : ['$third\_val', '$even\_third']

}

}

}

])