

DATA ANALYZATION AND VISUALISATION USING MONGODB SHELL AND MONGODB CHARTS

By G Snigda

Index

- Objective
- About the Dataset
- Analyzation
- Working on MongoDB Charts
- Demo of MongoDB Charts
- Dashboard of MongoDB Charts
- References

OBJECTIVE

- The objective of this project is to analyze the data using MongoDB Shell and also visualize using the MongoDB Charts. The dataset used in the project is ‘Campus Recruitment’ from Kaggle.
- This project mainly focuses on understanding how the academic and employability factors influence placements.
- In this project, we used MongoDB shell for analyzation which is a Document oriented NoSQL Database. Document consists of key-value pairs which are the basic unit of data in MongoDB.
- We also used MongoDB Charts for visualization of the data which a tool to create visual representations of the MongoDB data. Data visualization is a key component in providing a clear understanding of the data.

About the Dataset

- This dataset is a survey on the Campus Recruitments of 215 students of a college.
- This data set consists of Placement data of students in a campus. It includes secondary and higher secondary school percentage and specialization. It also includes degree specialization, type and Work experience and salary offers of the placed students.
- Link of the Dataset : <https://www.kaggle.com/benroshan/factors-affecting-campus-placement>

Analyzation

1. Importing the dataset(csv file) into MongoDB with database name as project and the collection name as data

```
Snigdas-MacBook-Air:~ gedelasnigda$ mongoimport --db project --collection data --type csv --headerline --ignoreBlanks --file /Users/gedelasnigda/Desktop/placements.csv
2020-06-08T02:18:05.967+0530      connected to: mongodb://localhost/
2020-06-08T02:18:06.076+0530      215 document(s) imported successfully. 0 document(s) failed to import.
Snigdas-MacBook-Air:~ gedelasnigda$ mongo
MongoDB shell version v4.2.7
connecting to: mongodb://127.0.0.1:27017/?compressors=disabled&gssapiServiceName=mongodb
Implicit session: session { "id" : UUID("311d11f9-5f91-4036-8199-d6bfff01d065") }
MongoDB server version: 4.2.7
Server has startup warnings:
2020-06-06T19:19:50.007+0530 I CONTROL  [initandlisten]
2020-06-06T19:19:50.010+0530 I CONTROL  [initandlisten] ** WARNING: Access control is not enabled for the
database.
2020-06-06T19:19:50.011+0530 I CONTROL  [initandlisten] **          Read and write access to data and conf
iguration is unrestricted.
2020-06-06T19:19:50.013+0530 I CONTROL  [initandlisten]
2020-06-06T19:19:50.015+0530 I CONTROL  [initandlisten]
2020-06-06T19:19:50.016+0530 I CONTROL  [initandlisten] ** WARNING: soft rlimits too low. Number of files
is 256, should be at least 1000
---
Enable MongoDB's free cloud-based monitoring service, which will then receive and display
metrics about your deployment (disk utilization, CPU, operation statistics, etc).

The monitoring data will be available on a MongoDB website with a unique URL accessible to you
and anyone you share the URL with. MongoDB may use this information to make product
improvements and to suggest MongoDB products and deployment options to you.

To enable free monitoring, run the following command: db.enableFreeMonitoring()
To permanently disable this reminder, run the following command: db.disableFreeMonitoring()
---]
```

2. Printing the first document of the dataset

This dataset has 15 attributes namely –

1. sl_no - serial No,
2. Gender - M or F
3. ssc_p - SSC Percentage
4. ssc_b – SSC Board of education
5. hsc_p – higher secondary Edu. Percentage
6. hsc_b - HSE Board of education
7. hsc_s – specialisation in HSE
8. degree_p – degree percentage
9. degree_t – under graduate degree type
10. workex – work experience (Yes or No)
11. etest_p – employability test percentage
12. sppecialisation – post graduation specialization(MBA)
13. mba_p – MBA percentage
14. status – status of placement (placed/not placed)
15. salary – salary offered by corporate to candidate

```
> show dbs
admin      0.000GB
config     0.000GB
local      0.000GB
projct     0.000GB
project    0.000GB
project1   0.000GB
> use project
switched to db project
> use project
switched to db project
> db.data.findOne()
{
  "_id" : ObjectId("5edd52853e8d9c4a5f6fe93d"),
  "sl_no" : 1,
  "gender" : "M",
  "ssc_p" : 67,
  "ssc_b" : "Others",
  "hsc_p" : 91,
  "hsc_b" : "Others",
  "hsc_s" : "Commerce",
  "degree_p" : 58,
  "degree_t" : "Sci&Tech",
  "workex" : "No",
  "etest_p" : 55,
  "specialisation" : "Mkt&HR",
  "mba_p" : 58.8,
  "status" : "Placed",
  "salary" : 270000
}
> █
```

3. Average salary of all the students

```
> db.data.aggregate([{"$match": {"status: "Placed" } }, {"$group": {"_id: null, avgSalary: { $avg: "$salary" }} }]).pretty()
{ "_id" : null, "avgSalary" : 288655.4054054054 }
>
```

Average Salary is 288655

4. Number of students who got salary above average

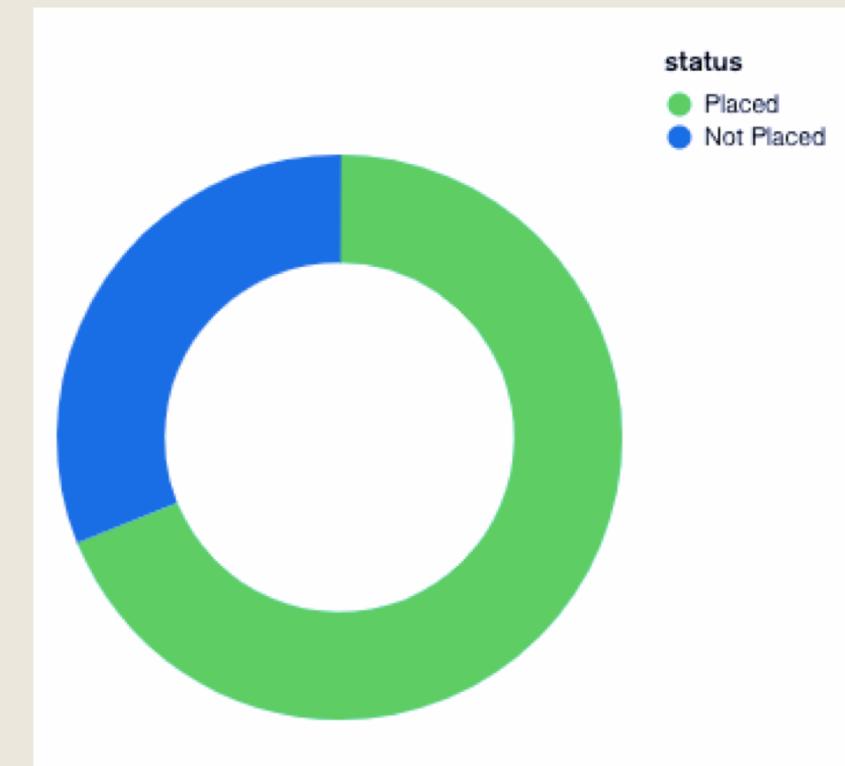
```
> db.data.aggregate([{"$match": { salary: {$gt:288655} } }, {"$count:"CountAbvAvg"}]).pretty()
{ "CountAbvAvg" : 52 }
>
```

52(out of 215) students got placed with salary above average.

5. Number of students who got placed

```
> db.data.aggregate([{"$match": { status:"Placed" } }, {"$count:"GotPlaced"}]).pretty()
{ "GotPlaced" : 148 }
>
```

148(out of 215) students got placed in total.

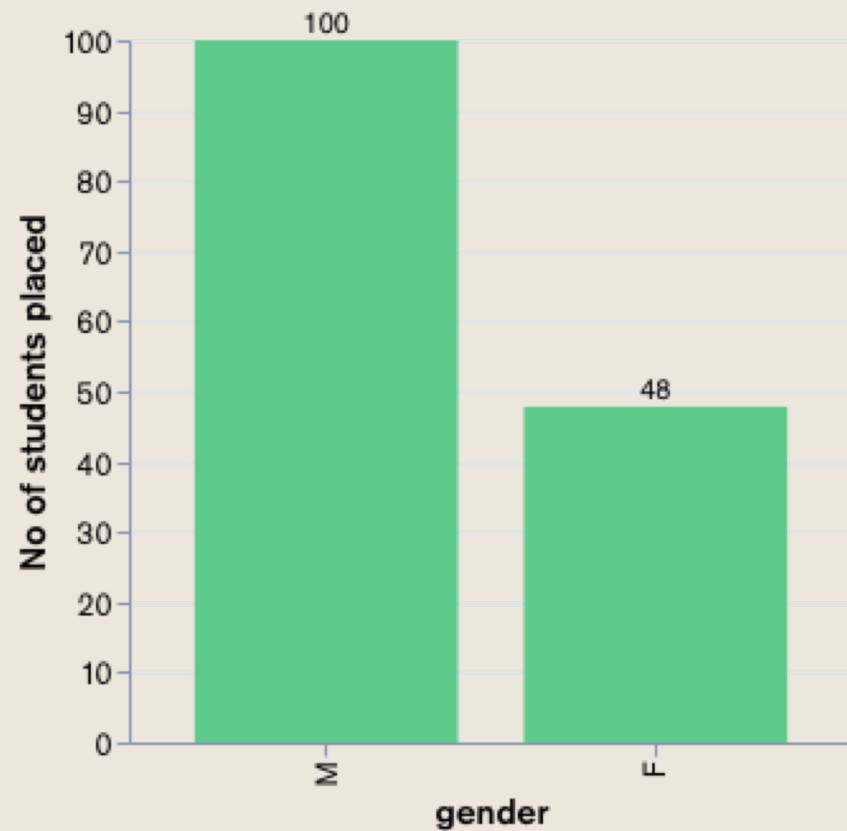


Here , it is seen that 148 students got placed out of 215 students in which 52 students got placed with a salary above the average salary(24%). This shows that overall the numbers of students who got placed is significant but the amount of salary they got offered by the corporate is not significant enough as only 24% of students got above average salary.

6. Number of males who got placed

It is seen from the mongo query and the chart that 100(out of 148 placed student) males got placed out of all the students who got placed. This shows that majority of the students who are getting well educated are still males.

```
> db.data.aggregate( [ { $match: { $and: [ { status:"Placed" }, { gender:"M" } ] } },  
{ $group: { _id: null, MalesGotPlaced: { $sum: 1 } } } ] ).pretty()  
{ "_id" : null, "MalesGotPlaced" : 100 }
```

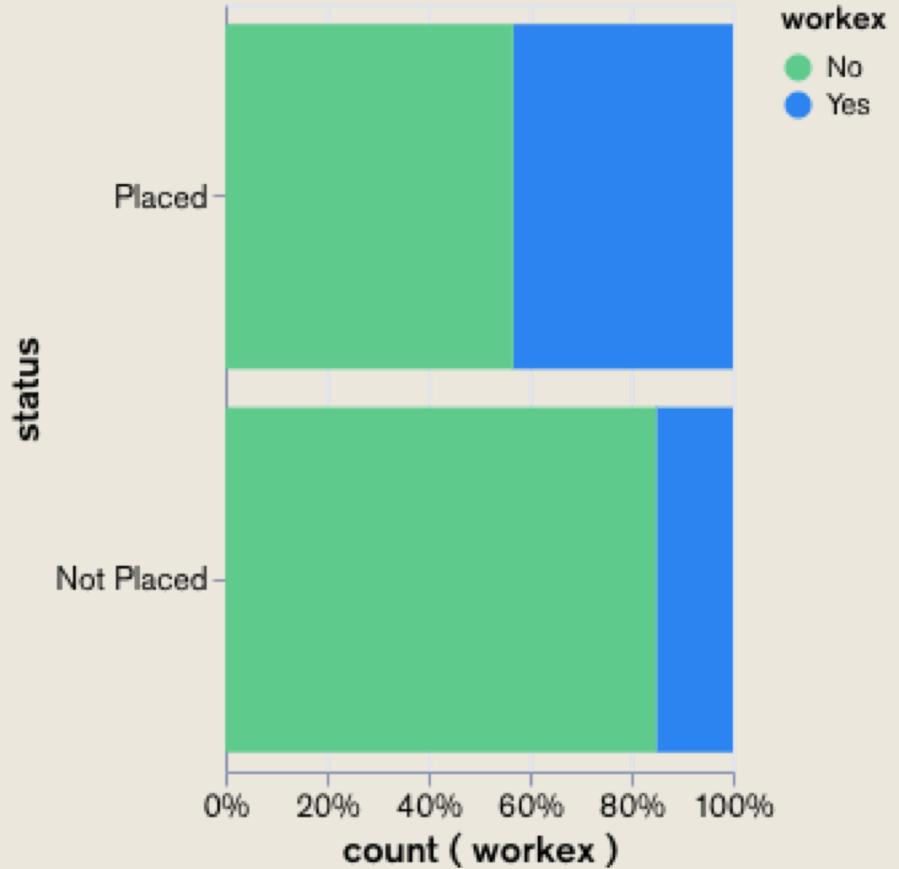


7. Does working experience matter for getting placed?

Only 64(out of 148 placed student) students have work experience. This shows that less than 50% of the students who got placed have work experience.

But the graph clearly indicates that more than 80% of the students who did not get placed do not have work experience.

Hence, We can conclude that work experience doesn't effect the chances of a student to get recruited.



```
> db.data.aggregate( [ { $match: { $and: [ { status:"Placed" }, { workex:"Yes" } ] } }, { $group: { _id: null, MalesGotPlaced: { $sum: 1 } } } ] ).pretty()
{ "_id" : null, "MalesGotPlaced" : 64 }
>
```

8. Do the employability tests help the students in getting placed?



8. Do the employability tests help the students in getting placed?

The analysis shows that 108 students(out of 215) got 70% above in etests out of which 79 students got placed.

The above graph shows that the ‘Not Placed’ line is gradually decreasing as the etest percentage increases.

Hence, we can conclude that etests do help the students in getting placed.

- Salary of top 5 etest scored students

```
> db.data.find({}, {status:1,salary:1}).sort({"etest_p":-1}).limit(5).pretty()
{
    "_id" : ObjectId("5edd52853e8d9c4a5f6fe9d5"),
    "status" : "Placed",
    "salary" : 240000
}
{
    "_id" : ObjectId("5edd52853e8d9c4a5f6fe955"),
    "status" : "Placed",
    "salary" : 360000
}
{
    "_id" : ObjectId("5edd52853e8d9c4a5f6fe9b3"),
    "status" : "Placed",
    "salary" : 276000
}
{
    "_id" : ObjectId("5edd52853e8d9c4a5f6fea0a"), "status" : "Not Placed"
}
{
    "_id" : ObjectId("5edd52853e8d9c4a5f6fe93e"),
    "status" : "Placed",
    "salary" : 425000
}
```

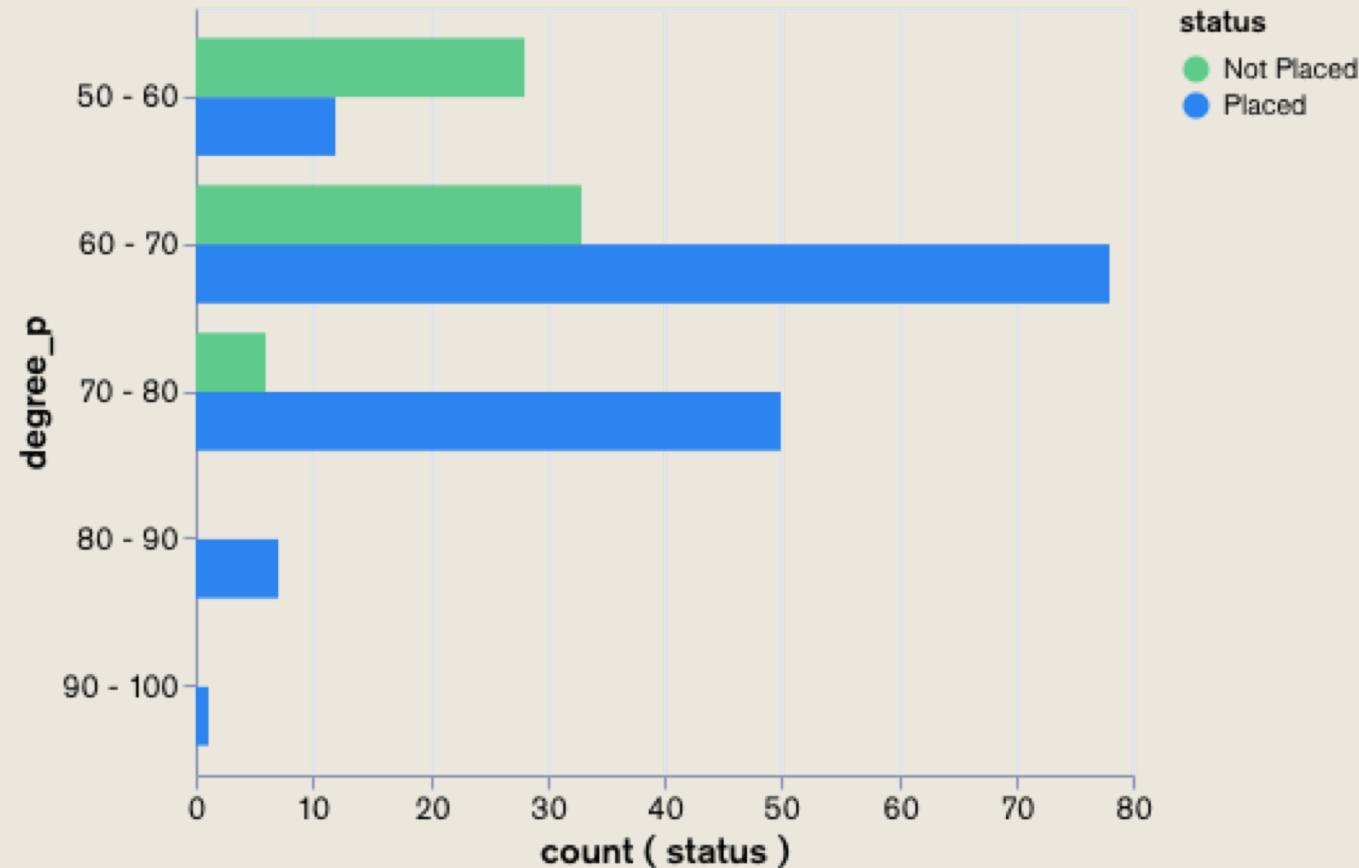
- Number of students who got etest more than 70%

```
> db.data.aggregate([{$match: { etest_p:{ $gt:70} } },{$group: { _id: null, total: { $sum:1 } } }])
{
    "_id" : null, "total" : 108
}
```

- Number of students who got etest more than 70% and got placed

```
> db.data.aggregate( [{ $match: { $and: [ {status:"Placed" }, { etest_p:{$gt:70} } ] } },{$group: { _id: null, count: { $sum: 1 } } }]).pretty()
{
    "_id" : null, "count" : 79
}
```

9. Does degree percentage matter for one to get placed?



9. Does degree percentage matter for one to get placed?

Among the first five students ranked according the salary, there is one case where the student got less than 60 percent in her degree. But according to the further analysis done, as the degree degree percentage limit increased, the percentage of the students getting placed increased. On an average a student should get around 68% so that he has higher possibility of getting placed in a company

From the graph, it is understood that there will be less chances to get placed in a company with a degree percentage ranging 50-60 , there are high chances of getting placed if the percentage is ranging 60-80 and will surely get placed if he has a percentage above 80.

Hence, we can say that degree percentage does matter for one to get placed.

- Percentage in Degree of the students who got highest salary(top 5)

```
[> db.data.find({},{degree_p:1}).sort({"salary":-1}).limit(5).pretty()
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe9b4"), "degree_p" : 64.6 }
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe9d3"), "degree_p" : 58 }
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe9ee"), "degree_p" : 79 }
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe98a"), "degree_p" : 65 }
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe9e0"), "degree_p" : 64 }
```

- Comparing student's degree_p and status

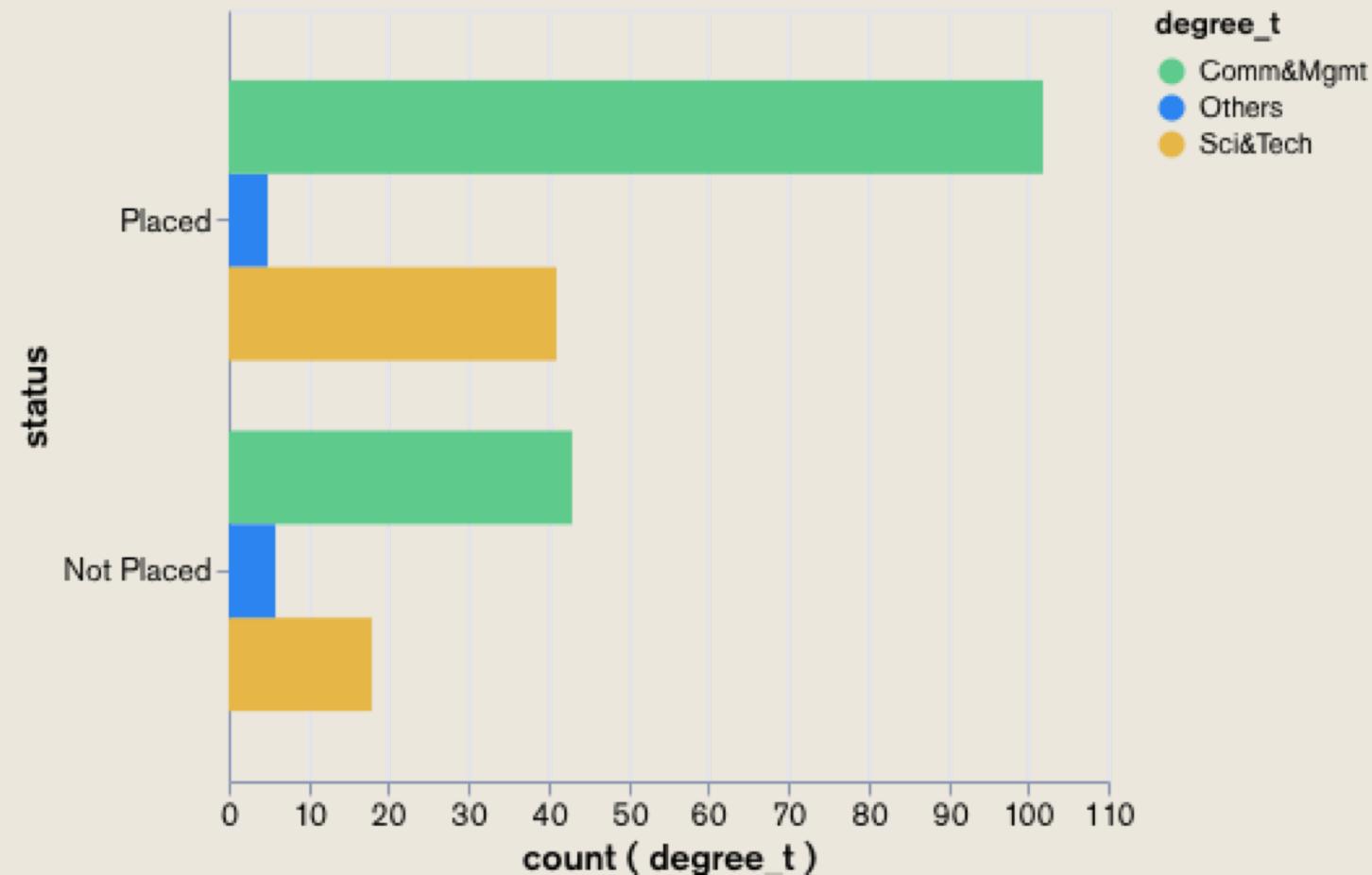
```
[> db.data.aggregate([{"$match": { degree_p:{ $gt:50} } }, {"$group": { _id: null, total:[{ $sum:1 } ]}}]
{ "_id" : null, "total" : 214 }
[> db.data.aggregate( [ { $match: { $and: [ {status:"Placed" }, { degree_p:{$gt:50} } ]} }, {"$group": { _id: null, count: { $sum: 1 } }}] ).pretty()
{ "_id" : null, "count" : 148 }
[> db.data.aggregate([{"$match": { degree_p:{ $gt:70} } }, {"$group": { _id: null, total:[{ $sum:1 } ]}}]
{ "_id" : null, "total" : 62 }
[> db.data.aggregate( [ { $match: { $and: [ {status:"Placed" }, { degree_p:{$gt:70} } ]} }, {"$group": { _id: null, count: { $sum: 1 } }}] ).pretty()
{ "_id" : null, "count" : 56 }
[>
[> db.data.aggregate([{"$match": { degree_p:{ $gt:60} } }, {"$group": { _id: null, total:[{ $sum:1 } ]}}]
{ "_id" : null, "total" : 167 }
[> db.data.aggregate( [ { $match: { $and: [ {status:"Placed" }, { degree_p:{$gt:60} } ]} }, {"$group": { _id: null, count: { $sum: 1 } }}] ).pretty()
{ "_id" : null, "count" : 133 }
```

x	50	60	70
degree_p above x	214	167	62
degree_p above x and placed	148	133	56
Percentage	69%	80%	90%

- Average percentage of the students who got placed

```
> db.data.aggregate( [{"$match: { status:"Placed" }}, {"$group: { _id: null, count: { $avg: "$degree_p" } }}] ).pretty()
{ "_id" : null, "count" : 68.74054054054054 }
```

10. Which degree is much demanded by corporate?



10. Which degree is much demanded by corporate?

In terms of the number of students placed in each degree, 70% of the Comm&Mgmt students got placed and 74% of the Sci&Tech students got placed. Even the graph clearly states that large number of Comm&Mgmt got placed but in terms of the average salary, the average salary of the Sci&Tech students is more than average salary of Comm&Mgmt students with only 55 students(90 students less than Comm&Mgmt students).

Hence, Sci&Tech degree has more demand than Comm&Mgmt by corporate because the students of Sci&Tech are getting paid highly.

- Number of students placed(in each degree)

```
[> db.data.aggregate([{"$match": {"degree_t": "Comm&Mgmt"}}, {"$group": {"_id": null, "total": {"$sum": 1}}}] )
{ "_id" : null, "total" : 145 }
[> db.data.aggregate( [{"$match": {"$and": [ {"status": "Placed"}, {"degree_t": "Comm&Mgmt"} ] }}, {"$group": {"_id": null, "count": {"$sum": 1}}}] ).pretty()
{ "_id" : null, "count" : 102 }
[> db.data.aggregate( [{"$match": {"$and": [ {"status": "Placed"}, {"degree_t": "Sci&Tech"} ] }}, {"$group": {"_id": null, "count": {"$sum": 1}}}] ).pretty()
{ "_id" : null, "count" : 41 }
```

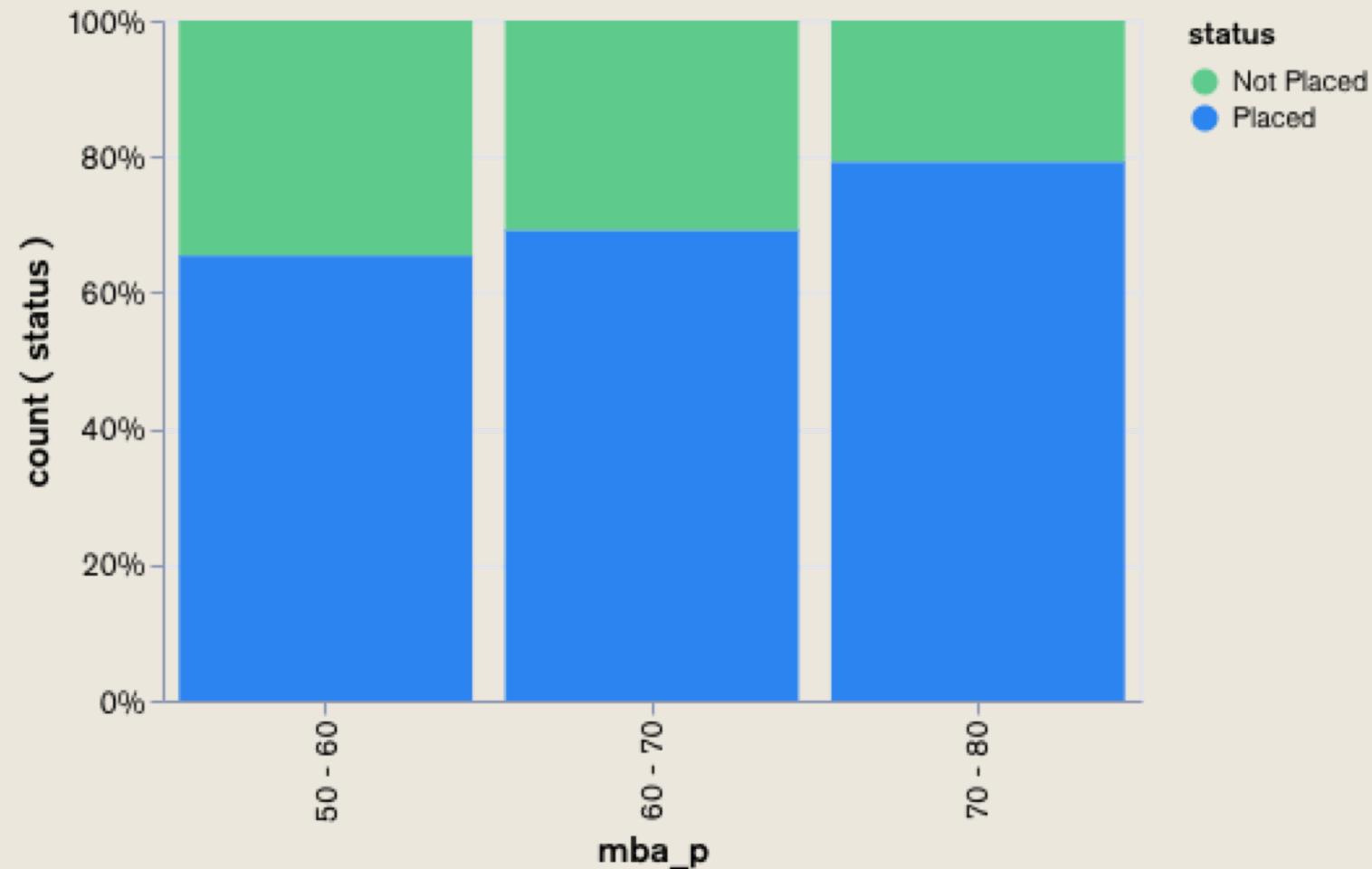
Total no of Comm&Mgmt students placed	102/145(70%)
Total no of Sci&Tech students placed	41/55(74%)

- Average salary of students of each degree

```
[> db.data.aggregate( [{"$match": {"$and": [ {"status": "Placed"}, {"degree_t": "Comm&Mgmt"} ] }}, {"$group": {"_id": null, "count": {"$avg": "$salary"}}}] ).pretty()
{ "_id" : null, "count" : 278627.45098039217 }
[> db.data.aggregate( [{"$match": {"$and": [ {"status": "Placed"}, {"degree_t": "Sci&Tech"} ] }}, {"$group": {"_id": null, "count": {"$avg": "$salary"}}}] ).pretty()
{ "_id" : null, "count" : 314609.756097561 }
```

Average salary of Comm&Mgmt students	278627
Average salary of Sci&Tech students	314609

11. Does MBA percentage matter for one to get placed?



11. Does MBA percentage matter for one to get placed?

Among the first five students ranked according to the salary, it is observed that all their MBA percentages range between 50-65%.

According to the comparison by the MBA percentage and the number of students places and looking at the graph it is is observed that there is no big change in the ratio of number of students getting placed.

So we can conclude that the possibility of getting placed in a company will increase by the increase in the MBA percentage but a significant increase wouldn't be visible. But a student has to get 62% on an average so that the possibility of getting placed increases.

- Percentage in MBA of the students who got highest salary(top 5)

```
> db.data.find({}, {mba_p:1}).limit(5).pretty()
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe93d"), "mba_p" : 58.8 }
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe93e"), "mba_p" : 55.5 }
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe93f"), "mba_p" : 66.28 }
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe940"), "mba_p" : 51.58 }
{ "_id" : ObjectId("5edd52853e8d9c4a5f6fe941"), "mba_p" : 53.29 }
```

- Comparing student's MBA percent and status

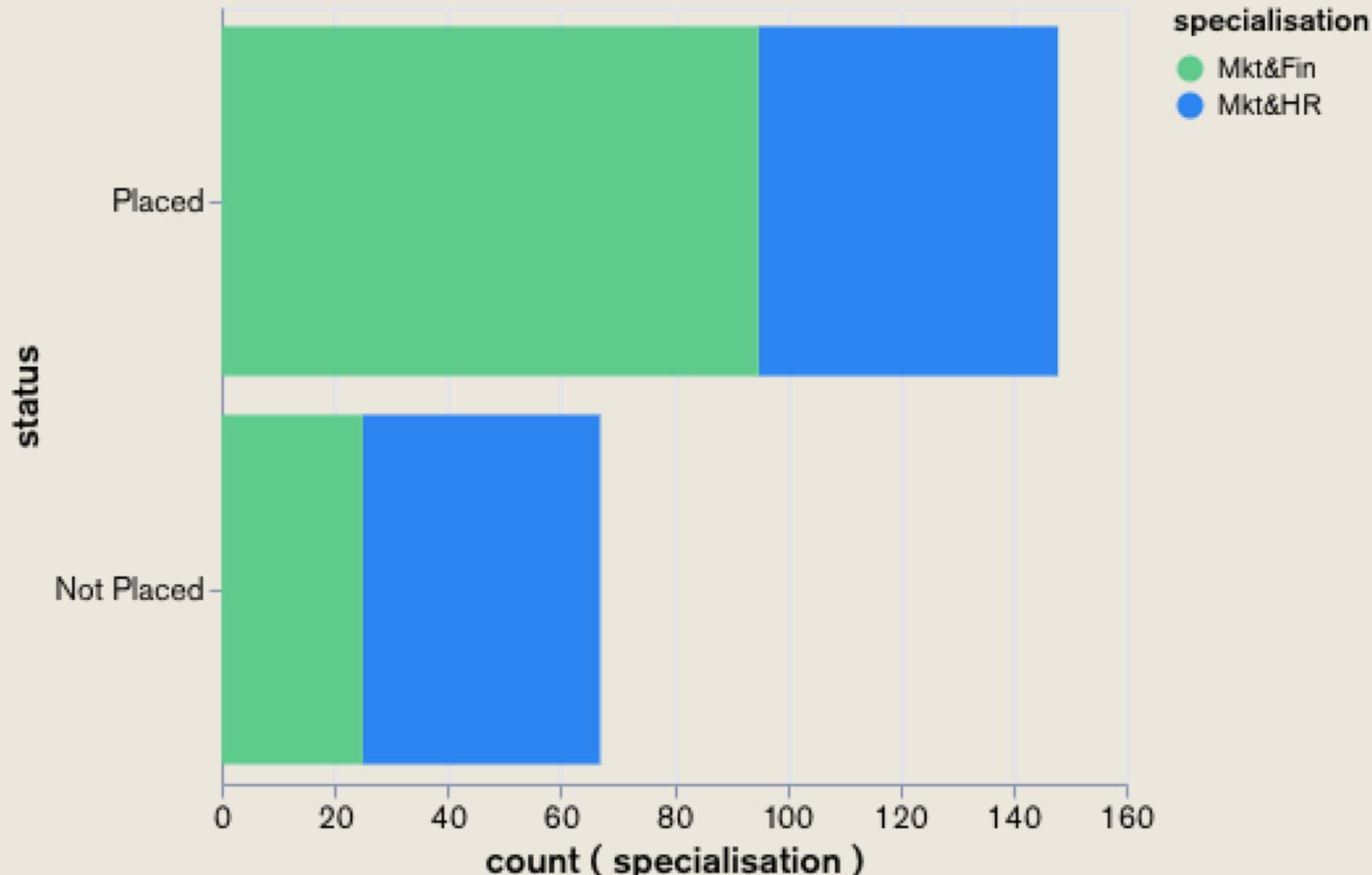
```
> db.data.aggregate([{$match: { mba_p:{ $gt:50} } },{$group: { _id: null, total: { $sum:1 } }}])
{ "_id" : null, "total" : 215 }
> db.data.aggregate( [{$match: { $and: [ {status:"Placed" }, { mba_p:{$gt:50} } ] }},{$group: { _id: null, count: { $sum: 1 } }}]).pretty()
{ "_id" : null, "count" : 148 }
> db.data.aggregate([{$match: { mba_p:{ $gt:60} } },{$group: { _id: null, total: { $sum:1 } }}])
{ "_id" : null, "total" : 134 }
> db.data.aggregate( [{$match: { $and: [ {status:"Placed" }, { mba_p:{$gt:60} } ] }},{$group: { _id: null, count: { $sum: 1 } }}]).pretty()
{ "_id" : null, "count" : 95 }
> db.data.aggregate([{$match: { mba_p:{ $gt:70} } },{$group: { _id: null, total: { $sum:1 } }}])
{ "_id" : null, "total" : 24 }
> db.data.aggregate( [{$match: { $and: [ {status:"Placed" }, { mba_p:{$gt:70} } ] }},{$group: { _id: null, count: { $sum: 1 } }}]).pretty()
{ "_id" : null, "count" : 19 }
```

x	50	60	70
degree_p above x	215	134	24
degree_p above x and placed	148	95	19
Percentage(%)	68	70	79

- Average percentage of the students who got placed

```
> db.data.aggregate( [{$match: { status:"Placed" }},{$group: { _id: null, count: { $avg: "$mba_p" } }}]).pretty()
{ "_id" : null, "count" : 62.579391891891895 }
```

12. Which degree specialization is much demanded by corporate?



12. Which degree specialization is much demanded by corporate?

In terms of the number of students placed in each degree specialisation, 79% of the Mkt&Fin students got placed and 66% of the Mkt&HR students got placed. In terms of the average salary, the average salary of the Mkt&Fin students is more than average salary of Mkt&HR.

Even by visualizing , in the above graph it is seen that more area of the bar is covered by the Mkt&Fn in case of 'Placed' and more area of the bar is covered by the Mkt&HR in case of 'Not Placed'.

Hence, we can conclude that Mkt&Fin is more demanded by corporate than the Mkt&HR degree specialization.

- Number of students placed(in each degree spec.)

```
> db.data.aggregate([{$match: {specialisation:"Mkt&Fin"}}, {$group: {_id: null, total: {$sum:1}} }])  
{ "_id" : null, "total" : 120 }  
> db.data.aggregate( [{ $match: { $and: [ {status:"Placed"}, { specialisation:"Mkt&Fin"} ] }},{ $group: {_id: null, count: {$sum: 1}} }]).pretty()  
{ "_id" : null, "count" : 95 }  
> db.data.aggregate( [{ $match: { $and: [ {status:"Placed"}, { specialisation:"Mkt&HR"} ] }},{ $group: {_id: null, count: {$sum: 1}} }]).pretty()  
{ "_id" : null, "count" : 53 }
```

Total no of Mkt&Fin students placed	95/120(79%)
Total no of Mkt&HR students placed	53/80(66%)

- Average salary of students each degree spec.

```
> db.data.aggregate( [{ $match: { $and: [ {status:"Placed"}, { specialisation:"Mkt&Fin"} ] }},{ $group: {_id: null, count: {$avg: "$salary"} } }]).pretty()  
{ "_id" : null, "count" : 298852.63157894736 }  
> db.data.aggregate( [{ $match: { $and: [ {status:"Placed"}, { specialisation:"Mkt&HR"} ] }},{ $group: {_id: null, count: {$avg: "$salary"} } }]).pretty()  
{ "_id" : null, "count" : 270377.358490566 }
```

Average salary of Mkt&Fin students	298852
Average salary of Mkt&HR students	270377

13. Does studying higher studies in a School of Central Board help in placements?

Placement of students who studied in sch...



Placement of students who did not study i...



13. Does studying higher studies in a School of Central Board help in placements?

According to the analysis, 57 students out of the 84 students who studied in a School of Central Board got placed and y results, 57 students out of the 84 students who studied in a School of Central Board got placed(i.e 67%) and 91 students out of the 131 students who did not study in a School of Central Board got placed(69%).

It is observed in the graphs above that in both the cases that is 'Student who studied in Central Board School' and 'Student who did not study in the Central Board School' the donut graphs look similar indicating that there is significant difference occurring in the placement recruitment due to the student's board of School.

Hence, we can conclude that studying higher studies in a school of Central Board does not affect the placement recruitment in any way.

```
[> db.data.aggregate( [ { $match: { hsc_b:"Central"} }, { $group: { _id: null, count: { $sum: 1 } } } ] ).pretty()
{ "_id" : null, "count" : 84 }
```

Number of students who studied in a School of Central Board : 84

```
[> db.data.aggregate( [ { $match: { $and: [ { status:"Placed" }, { hsc_b:"Central"} ] } }, { $group: { _id: null, count: { $sum: 1 } } } ] ).pretty()
{ "_id" : null, "count" : 57 }
```

Number of students who studied in a School of Central Board and got placed : 57

```
[> db.data.aggregate( [ { $match: { hsc_b:"Others"} }, { $group: { _id: null, count: { $sum: 1 } } } ] ).pretty()
{ "_id" : null, "count" : 131 }
```

Number of students who did not study in a School of Central Board : 131

```
[> db.data.aggregate( [ { $match: { $and: [ { status:"Placed" }, { hsc_b:"Others"} ] } }, { $group: { _id: null, count: { $sum: 1 } } } ] ).pretty()
{ "_id" : null, "count" : 91 }
```

Number of students who did not study in a School of Central Board and got placed : 91

Working in MongoDB Charts

- Importing the dataset into the Cluster of MongoDB Atlas.

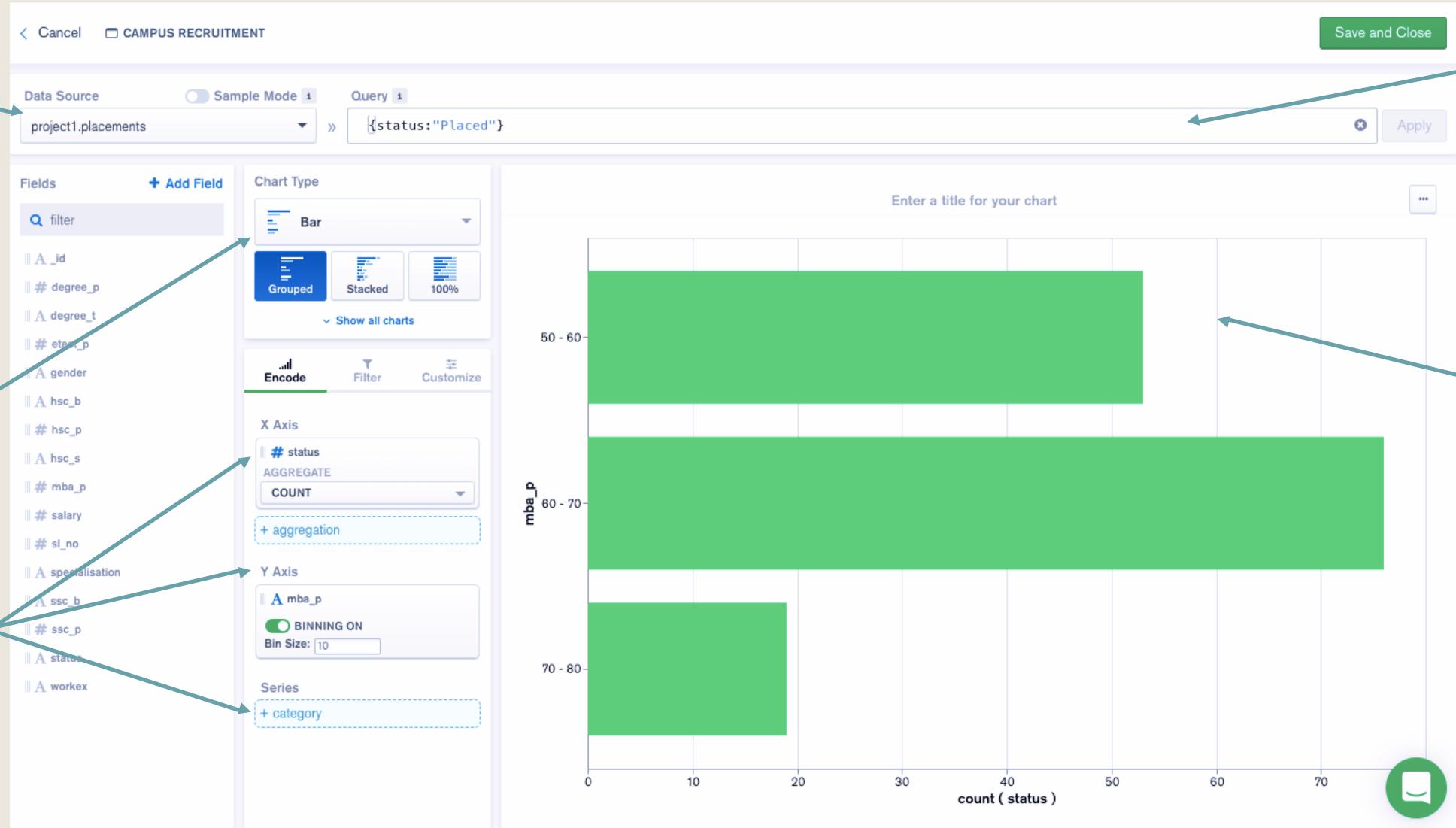
```
Snigdas-MacBook-Air:~ gedelasnigda$ mongoimport --host cluster0-shard-00-02-4fjgj.mongodb.net:27017  
--db project1 --collection placements --type csv --file /Users/gedelasnigda/Desktop/placements.csv -]  
-headerline --authenticationDatabase admin --ssl --username snigda0402 --password Tejovinay42  
2020-06-08T07:12:08.454+0530      connected to: mongodb://cluster0-shard-00-02-4fjgj.mongodb.net:27017]  
/  
2020-06-08T07:12:09.661+0530      215 document(s) imported successfully. 0 document(s) failed to impor  
t.
```

Demo of how I worked on MongoDB Charts

Dataset
is
chosen
here

Type of
graph is
chosen
here

X and Y
axis,
series
and
other
filters



MongoDB
Query is
applied
here.

The final
graph

Dashboard of MongoDB Charts



REFERENCES

- <https://www.youtube.com/watch?v=tpz-6Trd1UI>
- <https://charts.mongodb.com/charts-project-0-wslvb/dashboards/c84f73d0-d582-4db1-b8d4-301887f6f429/charts/create>
- <https://www.youtube.com/watch?v=yXp47VyXKWU>

THANK YOU