

The HR team is surfing social media to gather current and ex-employee feedback or sentiments. This information gathered will be used to derive actionable insights and take corrective actions to improve the employer-employee relationship. The data is web-scraped from Glassdoor and contains detailed reviews of 67K employees from Google, Amazon, Facebook, Apple, Microsoft, and Netflix.

```

[rm [-f] [-r] [-R] [-skipTrash] [-safely] <src> ...]
[rmr [-ignore-fail-on-non-empty] <dir> ...]
[setfacl [-R] [(b|k) [-m] [-x <acl_spec>] <path>][[--set <acl_spec>] <path>]
[setfacl [-n] [-m] [-v] <value>] [-x] <name>] <path>]
[setrep [-R] [-w] <rep> <path> ...]
[stat [-format] <path> ...]
[tail [-f] <file>]
[test [-dfs] <path>]
[test [-ignoreCrc] <src> ...]
[touch [-a] [-m] [-t] <TIMESTAMP>] [-c] <path> ...]
[touch <path> ...]
[truncate [-w] <length> <path> ...]
[usage [cmd ...]]

generic options supported are:
conf <configuration file> specify an application configuration file
D <property=value> define a value for a given property
fs <file:///hdfs://namenode:port> specify default filesystem URI to use, overrides 'fs.defaultFS' property from configurations.
jt <localResourceManager> specify a ResourceManager
files <file1...> specify a comma-separated list of files to be copied to the map reduce cluster
libjars <jar1,...> specify a comma-separated list of jar files to be included in the classpath
archives <archive1,...> specify a comma-separated list of archives to be unarchived on the compute machines

the general command line syntax is:
command [genericOptions] [commandOptions]

akshaykr951@gmail@ip-10-0-42-218 ~$ hdfs dfs -put akshay.txt
akshaykr951@gmail@ip-10-0-42-218 ~$ ls
kraj.txt akshay.txt akxy.txt test_account testsimplilearn.txt
akshaykr951@gmail@ip-10-0-42-218 ~$ ls
kraj.txt akshay.txt akxy.txt test_account testsimplilearn.txt
akshaykr951@gmail@ip-10-0-42-218 ~$ hdfs dfs -put akraj.txt Akshay123
akshaykr951@gmail@ip-10-0-42-218 ~$ hdfs dfs -get test_account
get: 'test_account': No such file or directory
akshaykr951@gmail@ip-10-0-42-218 ~$ hdfs dfs -get test_account
get: 'test_account/akshay_account': file exists
akshaykr951@gmail@ip-10-0-42-218 ~$ ls
kraj.txt akshay.txt akxy.txt test_account testsimplilearn.txt
akshaykr951@gmail@ip-10-0-42-218 ~$ cd test_account/
akshaykr951@gmail@ip-10-0-42-218 test_account$ ls
kshay_account employee_review_data.csv
akshaykr951@gmail@ip-10-0-42-218 test_account$ cd
akshaykr951@gmail@ip-10-0-42-218 ~$ ls
kraj.txt akshay.txt akxy.txt test_account testsimplilearn.txt
akshaykr951@gmail@ip-10-0-42-218 ~$ hdfs dfs -put employee_review_data.csv test_account
put: 'employee_review_data.csv': No such file or directory
akshaykr951@gmail@ip-10-0-42-218 ~$ cd test_account$ ls
kshay_account employee_review_data.csv
akshaykr951@gmail@ip-10-0-42-218 test_account$ hdfs dfs -put employee_review_data.csv test_account
akshaykr951@gmail@ip-10-0-42-218 test_account$
  
```

Hue - Editor | Chudra Manager | Chudra Manager | History Server | Learning on Simplilearn

sbdt-hue.coresstackio/hue/editor/editor=59259

Query: `select * from default.'employee_review_data' LIMIT 67539;`

0.70s Database default Type text

Tables (1800) +

Filter...

- 'temp_details'
- 'temp_review'
- 'thproject_step2'
- 'thproject_step20'
- 'temp_rating'
- '2temp_review'
- '2temp_review_25percent'
- '2temp_rating'
- '2temp_review'
- '2temp_review_25percent'
- '4temp_rating'
- '4temp_review'
- '4temp_review_25percent'
- '4temp_rating'
- '5_220k_wards_by_directors_case_study'
- '6_traffic_accidents_by_time_of_occurrence_'
- '6temp_review'
- '6temp_review_25percent'
- '6temp_rating'
- '17feb_stockcompanies'
- '_from_spark_stg_cars'
- 'aaa777'
- 'aag'
- 'abc12'
- 'abhi0410'
- 'abhi_001'
- 'abhi_data'
- 'abhi_parestock'
- 'abhi_parestock1'
- 'abhi_seq'
- 'abhi_stock'
- 'abhinav_stock_project'
- 'action'
- 'adash_prj_stock'

Query History | Saved Queries | Results (200+)

	employee_review_data.index	employee_review_data.company	employee_review_data.location	employee_review_data.dates	employee_review_data.job-title
1	1	google	none	NULL	Current Employee - Anonymous Employee
2	2	google	Mountain View, CA	NULL	Former Employee - Program Manager
3	3	google	New York, NY	NULL	Current Employee - Software Engineer III
4	4	google	Mountain View, CA	NULL	Current Employee - Anonymous Employee
5	5	google	Los Angeles, CA	NULL	Former Employee - Software Engineer
6	6	google	Mountain View, CA	NULL	Former Employee - SDE2
7	7	google	New York, NY	NULL	Current Employee - Software Engineer
8	8	google	none	NULL	Former Employee - Anonymous Employee
9	9	google	New York, NY	NULL	Current Employee - Anonymous Employee
10	10	google	none	NULL	Current Employee - Anonymous Employee
11	11	google	Mountain View, CA	NULL	Current Employee - Anonymous Employee
12	12	google	none	NULL	Current Employee - Anonymous Employee
13	13	google	none	NULL	Former Employee - Anonymous Employee
14	14	google	San Francisco, CA	NULL	Current Employee - Sales Operations Analyst
15	15	google	none	NULL	Current Employee - Anonymous Employee

0303 AM 27-04-2021

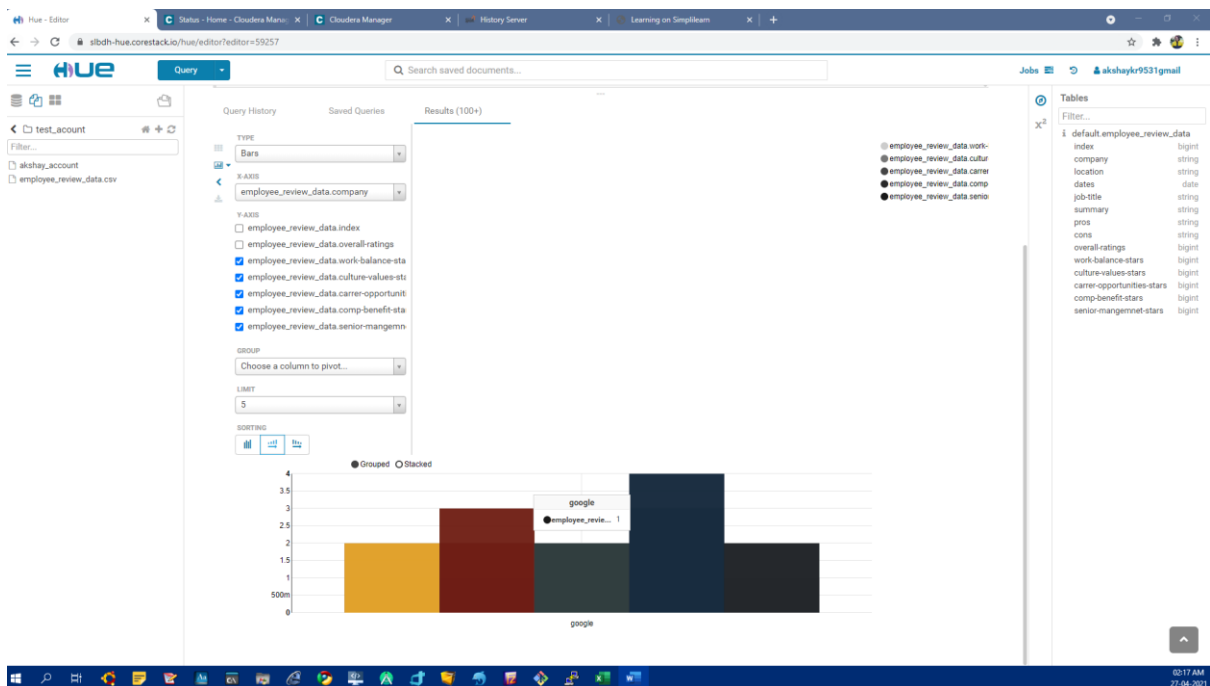
Query: `Select * From default.'employee_review_data' LIMIT 67538;`

Results (200+)

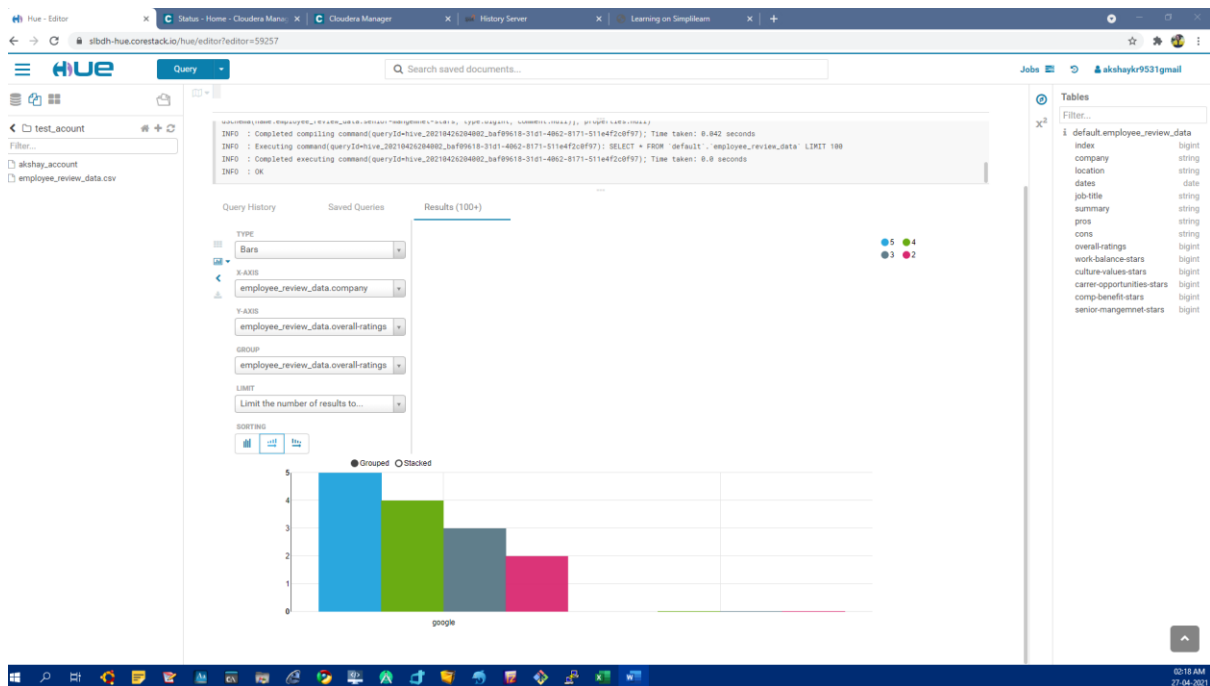
employee_review_data.index	employee_review_data.company	employee_review_data.location	employee_review_data.location	employee_review_data.location	employee_review_data.job-title
1	google	none	NULL	NULL	Current Employee - Anonymous Employee
2	google	Mountain View, CA	NULL	NULL	Former Employee - Program Manager
3	google	New York, NY	NULL	NULL	Current Employee - Software Engineer III
4	google	Mountain View, CA	NULL	NULL	Current Employee - Anonymous Employee
5	google	Los Angeles, CA	NULL	NULL	Former Employee - Software Engineer
6	google	Mountain View, CA	NULL	NULL	Former Employee - SDE2
7	google	New York, NY	NULL	NULL	Current Employee - Software Engineer
8	google	none	NULL	NULL	Former Employee - Anonymous Employee
9	google	New York, NY	NULL	NULL	Current Employee - Anonymous Employee
10	google	none	NULL	NULL	Current Employee - Anonymous Employee
11	google	Mountain View, CA	NULL	NULL	Current Employee - Anonymous Employee
12	google	none	NULL	NULL	Current Employee - Anonymous Employee
13	google	none	NULL	NULL	Former Employee - Anonymous Employee
14	google	San Francisco, CA	NULL	NULL	Current Employee - Sales Operations Analyst
15	google	none	NULL	NULL	Current Employee - Anonymous Employee

For imputation, calculate the median for each of the 5 rating fields and create a new table.

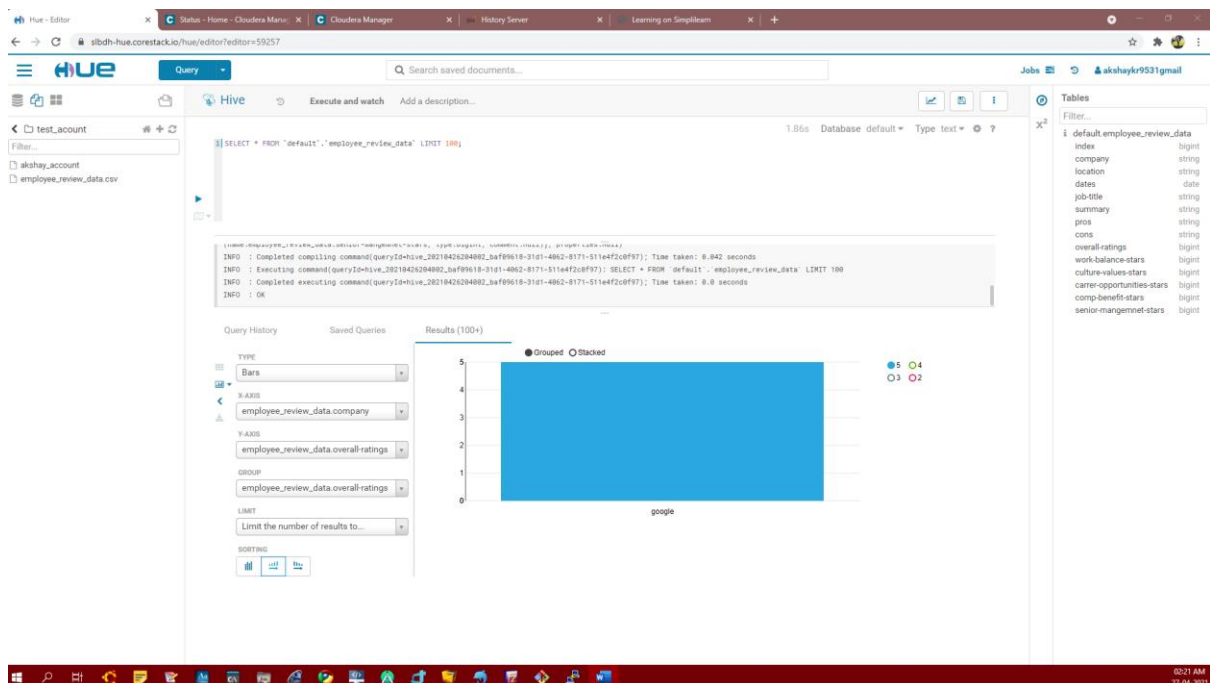
1. Work-balance stars
2. Culture values stars
3. Career opportunities-stars
4. Comp-benefit-stars
5. Senior-management-stars



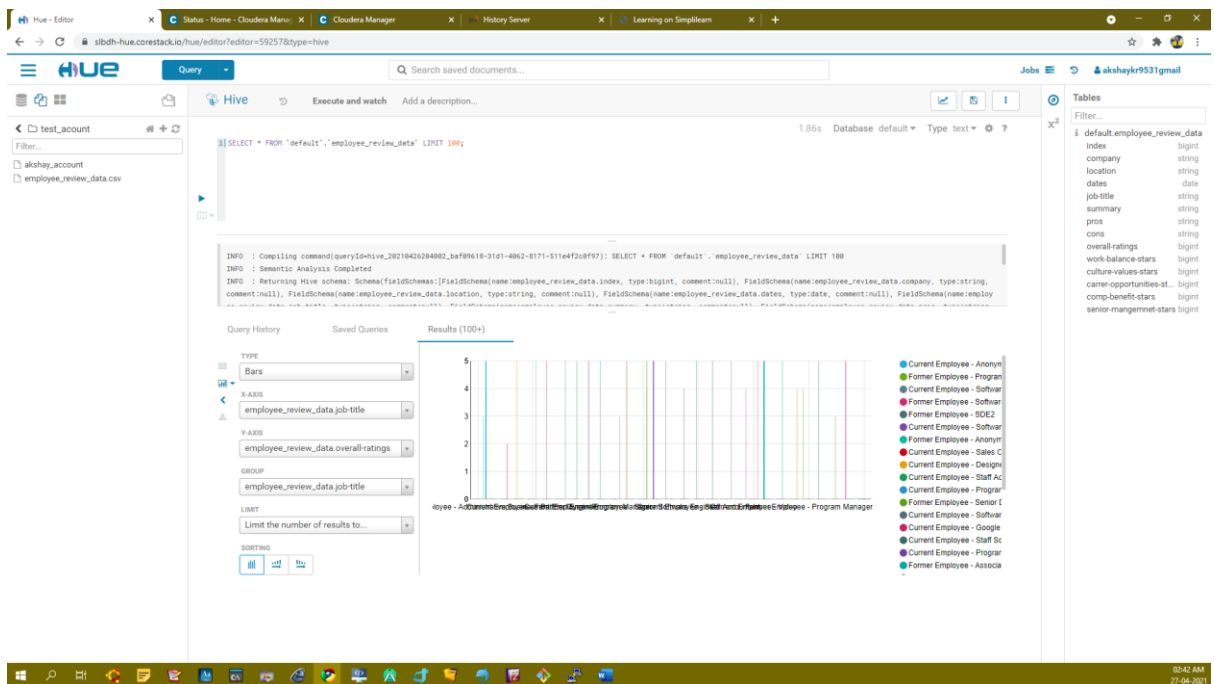
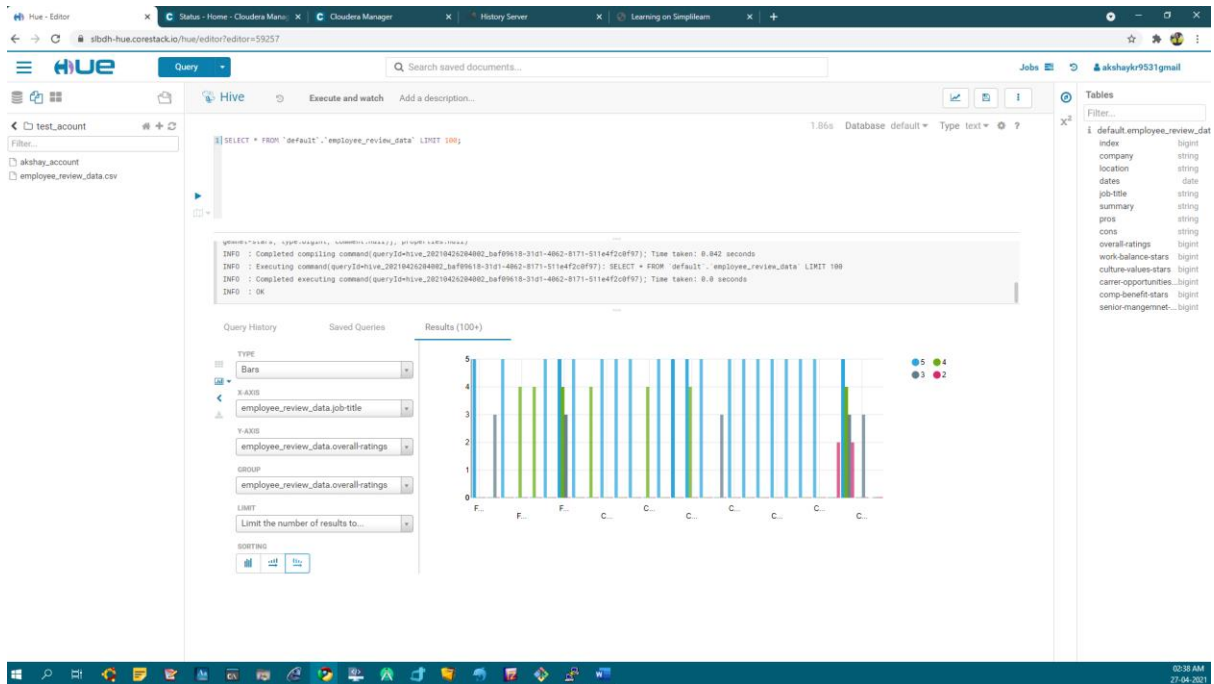
Graph BY Rating Over all :

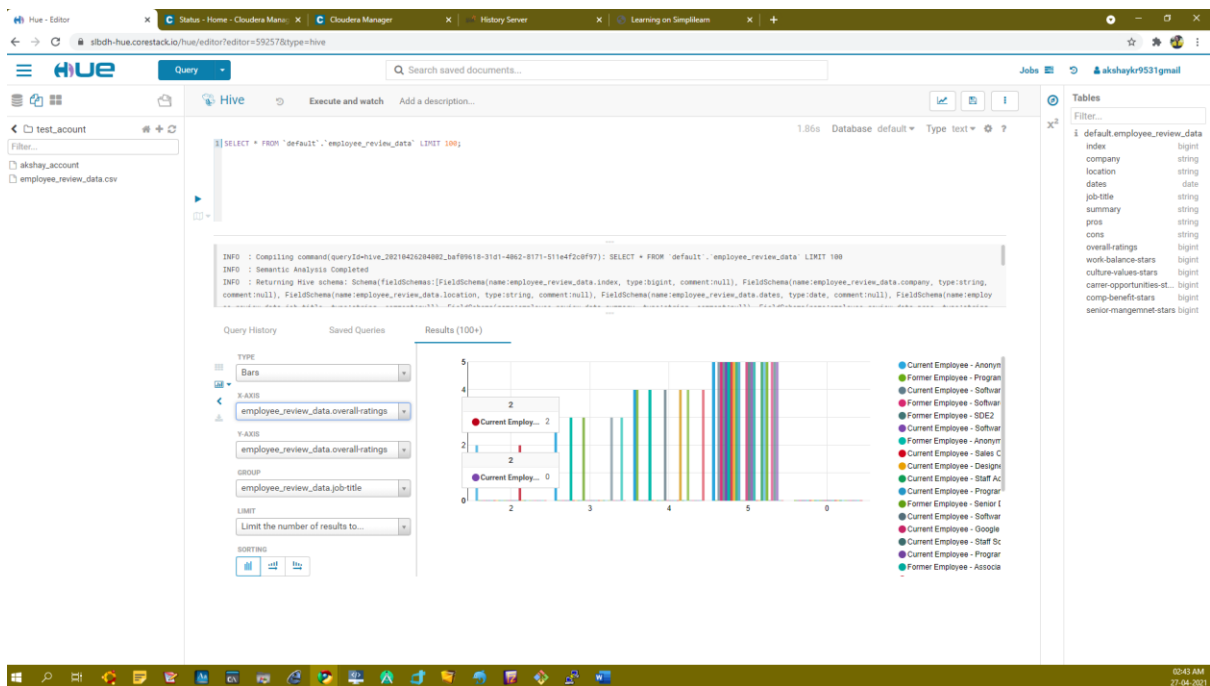


Graph by Rating only 5 :

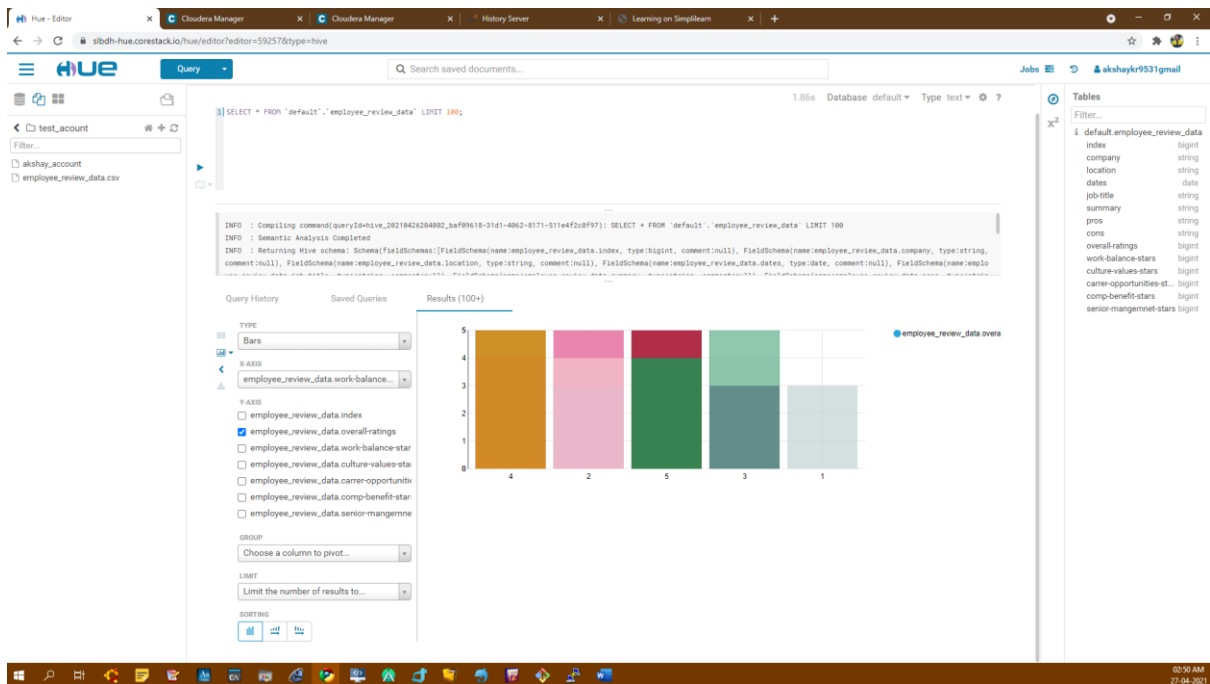


- Display the impact of job role on rating a company using the overall-ratings field by the company by year.

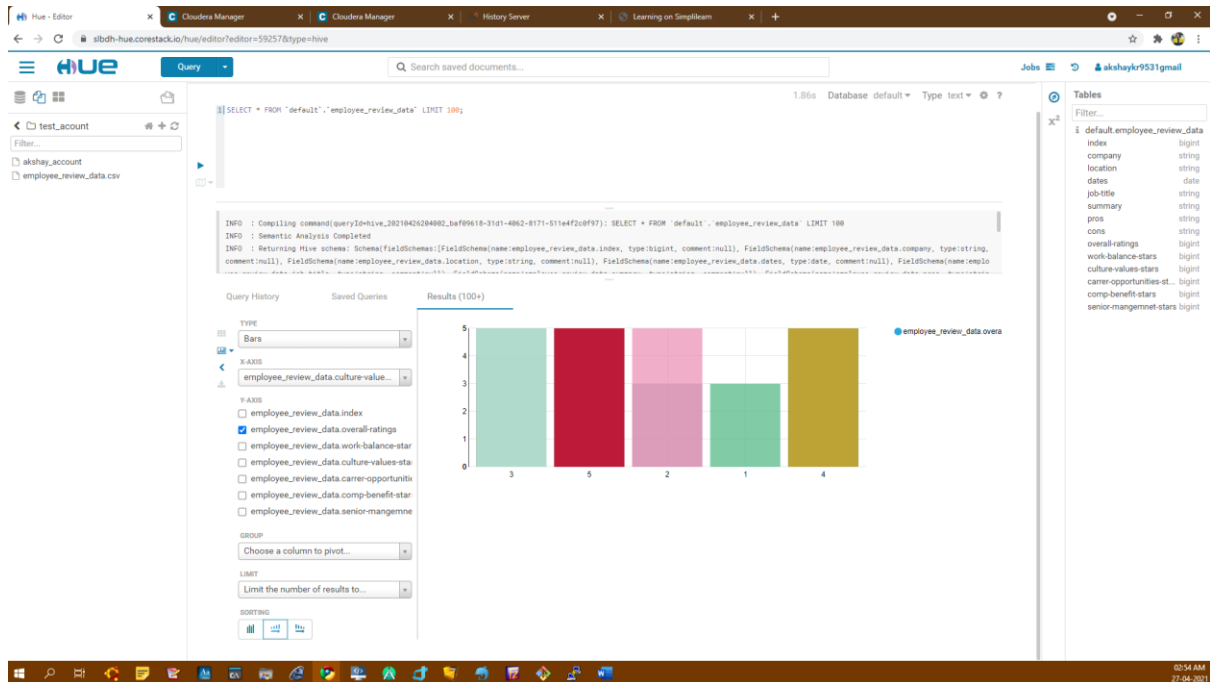




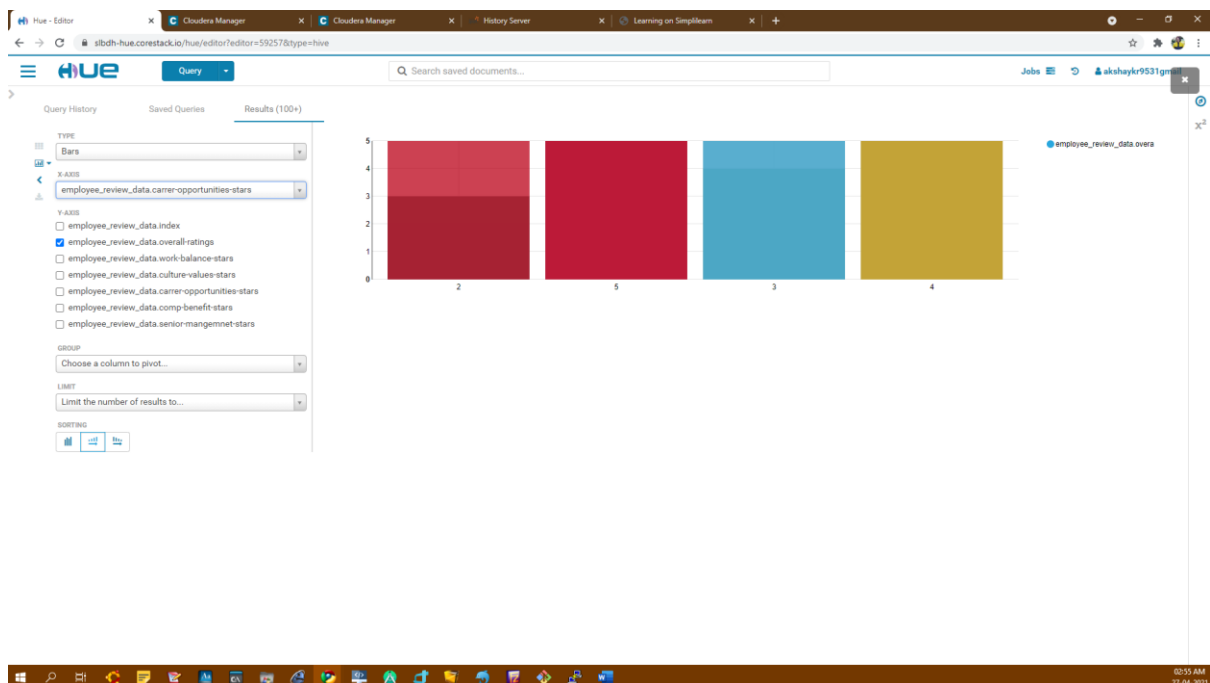
- Overall-ratings Versus
- Work-balance stars



Overall-ratings Versus Culture values stars



- Career opportunities-stars



- Comp-benefit-stars

