

Client Segmentation

PURPOSE

To perform quick segmentation on the client titles to help Marketing better target their campaigns.

Data Used: CSV file provided by Laura Brennan consisting of the following attributes:

- User id
- Company Id
- Client Title
- Gender
- Picture
- Role

Technology Used: Python

Quick Summary:

Laura Brennen signed off on the outcomes of the project which was mainly focused on coming up with client title segments in order to target marketing campaigns and tailor the language to make it specific to the client designations that they are targeting. The client title segments are intended to be added to the single source of truth repository (DOMO/Salesforce/Hubgroup). But as of Apr 9, 2018 a single source of truth has not been decided yet and the title segments have not been added to any source repository.

Update: Waiting for Marketing team to resolve bottlenecks with respect to deciding a single source of data repository.

A. Methods

There are currently 8144 client users that Shiftgig deals with on a daily basis. At present, there are 2133 client titles in Shiftgig's database. These are huge numbers that make it difficult for the marketing team to deliver right content to the right people in the right way. The main aim of the project was to come up with client title groups to which these 2133 titles could be mapped to so that the marketing team could target the right audience with appropriate campaigns. The entire approach was implemented using an SQL approach and the query was handed over to the BI team. The query can be found in the APPENDIX of this document.

I. CHALLENGES

The challenges that were encountered during the phase of the project were as follows:

- Misspelt words
E.g Manager was misspelled as manger. Administrator was misspelled as adminitrator
- Synonyms
E.g co-ordinator, planner, on-site all these titles almost meant the same
- Abbreviations
E.g whether to come up with specific segments for AM, Pm or use a general segments for all managers
- Reused Keywords
E.g executive keyword was part of executive chef and Chief Executive Officer each of which were to be assigned to different title segments

II. APPROACH

- **Binning**
 - After talks with Laura, the following strategic map containing common keywords was initially drafted.
- **Data Cleaning**
 - Punctuations, Apostrophe's and other irrelevant characters were dealt with the regular expression treatment after converting all the client titles to

lowercase. The regular expression treatment filled the irrelevant characters with spaces

- A pre-built python library was used to correct the misspelled words which were then added to the initial identification schema

- **Boolean Dummy Variable Approach**

- These 11 segments became variables which were used as an indicator of which bin a particular record falls into. Using these dummy variables, count of members in each client segment was retrieved and individual CSV files for each client segment was created.

- **Hierarchy**

- A hierarchical sorting was used to determine the precedence of the client titles.

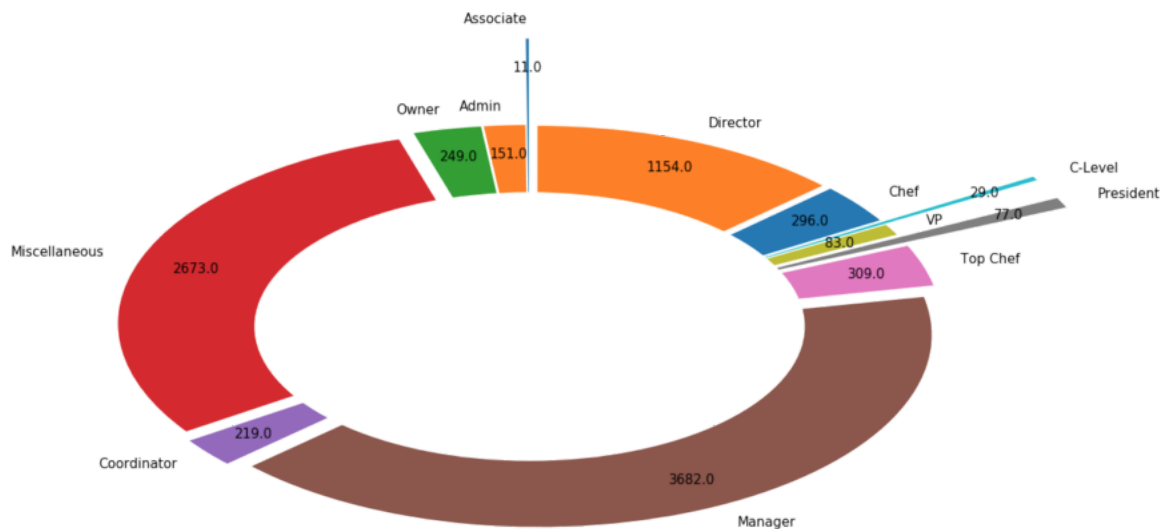
- **Valentina Postgres SQL query**

- The code in Python was replicated with a Postgres SQL approach and shared with the BI team. The query is added to the APPENDIX in the end of the report.

B. Results

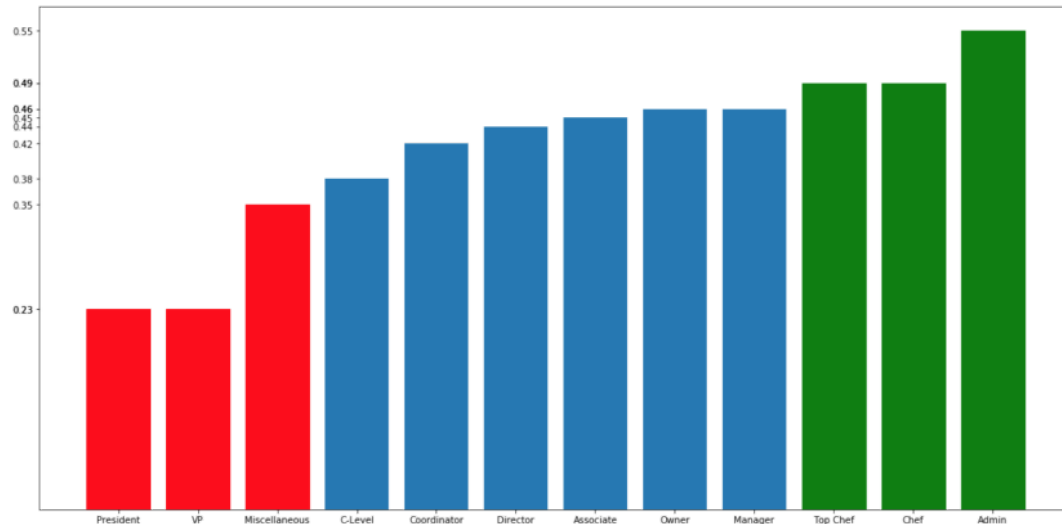
CSV files for the following individual client segments in descending order of ranking were created and sent out to the BI and Sales teams. One can find the files along with the SQL file [here](#).

It was interesting to look at the count of client users in each of the 11 client title segments. The following figure shows a pictorial representation of the same.



It was observed that major proportion of our client users are Managers. A significant chunk of client users are Directors, Chefs, so on and so forth. I was not able to cater a Miscellaneous client title segment. It consisted of titles like unknown, no longer with the company, demo acct, test profiles etc.

Considering an application of these segmented client users, I looked into the database to get an idea about the % of client users in each client title segment that have submitted an outcome. The statistics can be observed in the following figure.



It was seen that President, VP and Miscellaneous client title segments have the lowest % of client users that have submitted an outcome, whereas Manager, Chefs and Admin formed the segments that have the highest % of client users with a submitted outcome.

C. Conclusion

These results are still **preliminary** and may contain confounds or different approaches, still they help set benchmarks for the client segmentation project, define a language, and offer a hopeful perspective.

The approach still needs much improvement. This work helps set benchmarks and define the language on client segmentation but does not yet represent a productionalizable output. This work defined a foundational framework which could be used as a comparative base for future work using other approaches like text mining.

D. Future Directions

- Remove redundant client titles in Shiftgig's repository
- Come up with a centralized and consistent dataset that could be used for target marketing

E. APPENDIX

SELECT

id,
title,
CASE

WHEN (
title ILIKE '%owner%' OR
title ILIKE '%founder%'
) THEN 'Owner'

WHEN (
title ILIKE '%chef%'
AND (
title ILIKE '%exec%'
OR title ILIKE '%Head%'
OR title ILIKE '%ex.%'
OR title ILIKE '%Executive%'
)
) THEN 'Top Chef'

WHEN (
title ILIKE '%chef%'
OR TITLE ILIKE '%cook%'
) THEN 'Chef'

WHEN (
title ~* '\yVP\y' OR
title ~* '\ySVP\y' OR
title ~* '\yRVP\y' OR
title ~* '\yEVP\y' OR
(
title ILIKE '%Vice%'
AND title ILIKE '%President%'
)
) THEN 'VP'

WHEN (
title ILIKE '%President%' AND title NOT ILIKE '%Vice%'

```
OR title ~* '\yCEO\y'
OR title ILIKE '%Chief%'
AND (
    title ILIKE '%Executive%'
    AND (
        title ILIKE '%Officer%')
    )
) THEN 'President'
```

```
WHEN (
    title ILIKE '%Chief%'
    OR TITLE ~* '\yCTO\y'
    OR title ~* '\yCRO\y'
    OR title ~* '\yCFO\y'
    OR title ~* '\yCOO\y'
) THEN 'C-Level'
```

```
WHEN (
    title ILIKE '%director%'
    OR TITLE ILIKE '%controller%'
    OR title ILIKE '%dir%'
    OR title ILIKE '%MD%'
) THEN 'Director'
```

```
WHEN (
    title ILIKE '%Admin%'
    OR TITLE ILIKE '%administrator%'
    OR title ILIKE '%administrative%'
    OR title ILIKE '%analyst%'
) THEN 'Admin'
```

```
WHEN (
    title ILIKE '%associate%'
) THEN 'Associate'
```

```
WHEN (
    title ILIKE '%Coordinator%'
    OR title ILIKE '%planner%'
    OR title ILIKE '%on-site%'
    OR title ILIKE '%On Site%'
```

) THEN 'Coordinator'

WHEN (

title ILIKE '%Manager%'

OR title ILIKE '%manger%'

OR title ILIKE '%management%'

OR title ILIKE '%lead%'

OR title ILIKE '%supervisor%'

OR title ILIKE '%head%'

OR title ILIKE '%captain%'

OR title ILIKE '%mgr%'

OR title ILIKE '%Mgr%'

OR title ILIKE '%MGR%'

OR title ILIKE '%mngr%'

OR title ILIKE '%Mngr%'

OR title ILIKE '%gm%'

OR title ILIKE '%GM%'

OR title ILIKE '%Gm%'

OR title ILIKE '%agm%'

OR title ILIKE '%Agm%'

OR title ~* '\yPM\y'

OR title ~* '\yAM\y'

OR title ILIKE '%general%'

AND(

title ILIKE '%manager%'

)

) THEN 'Manager'

ELSE 'Misc.'

END AS inferred_title_level

FROM public.client