CIOs at UVA

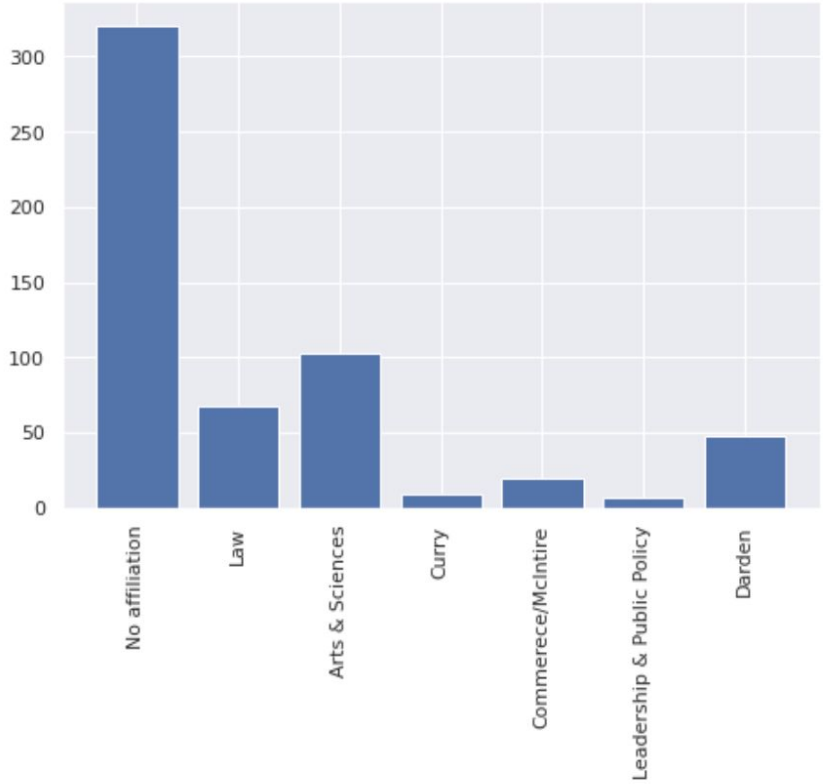
Clown Computing © Est. 2018 L.L.C

Introduction

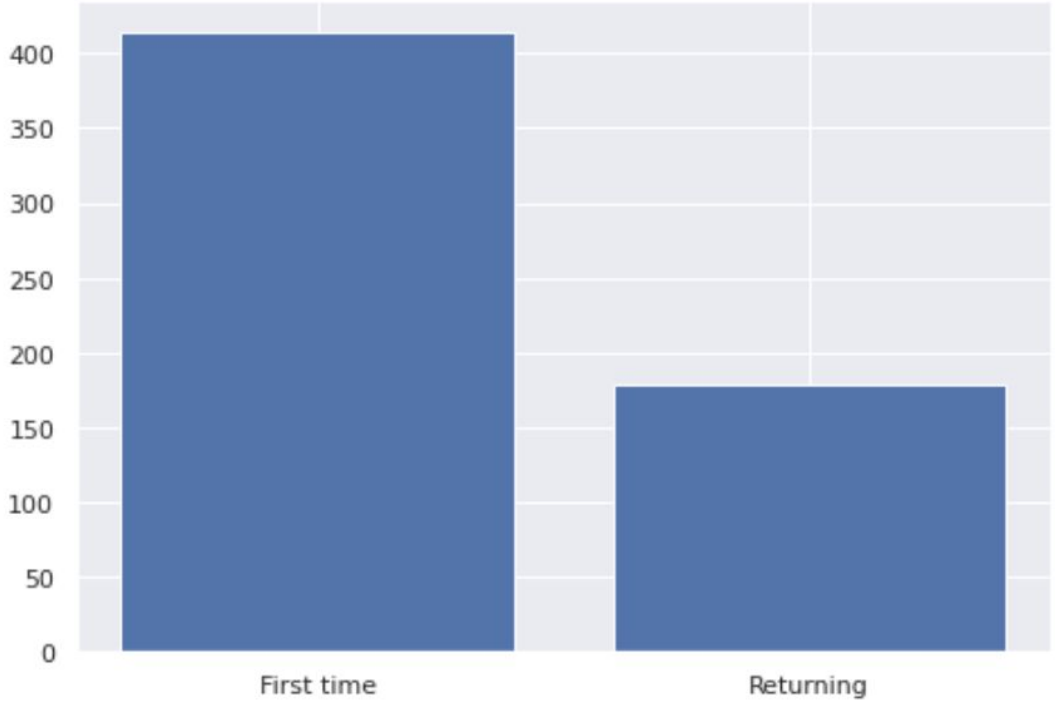
Goal: Accurately **measure** and **predict** the engagement levels of UVa students within CIOs based on given student and CIO attributes.

Data: Consisted of 34 columns, and 644 observations

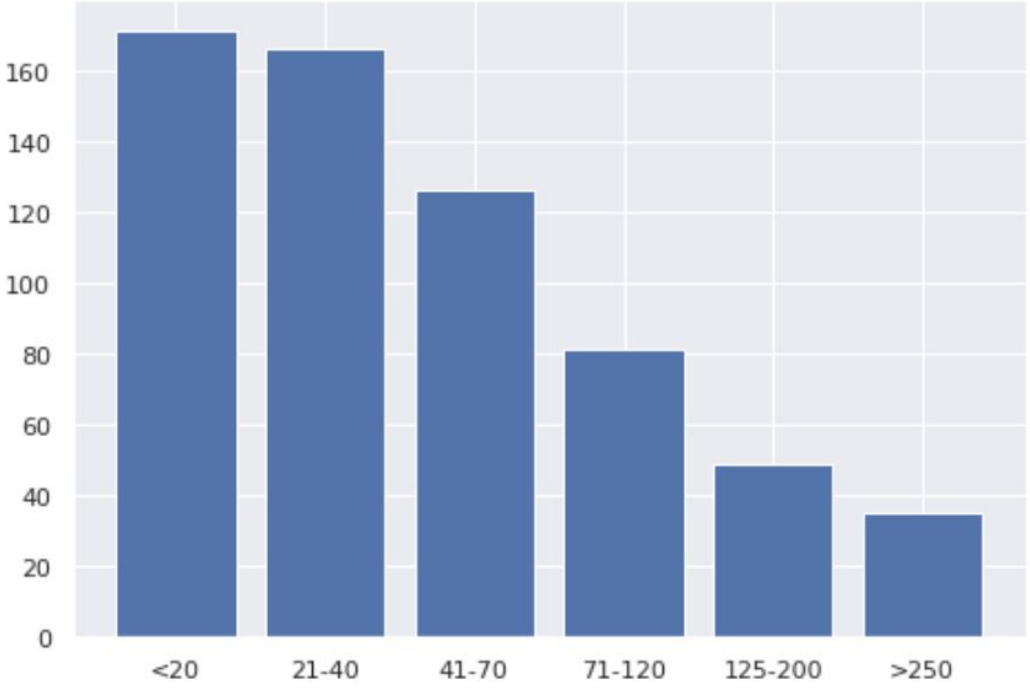
Affiliated School Distribution



First Time President Count



Distribution of Total Members



Cleaning

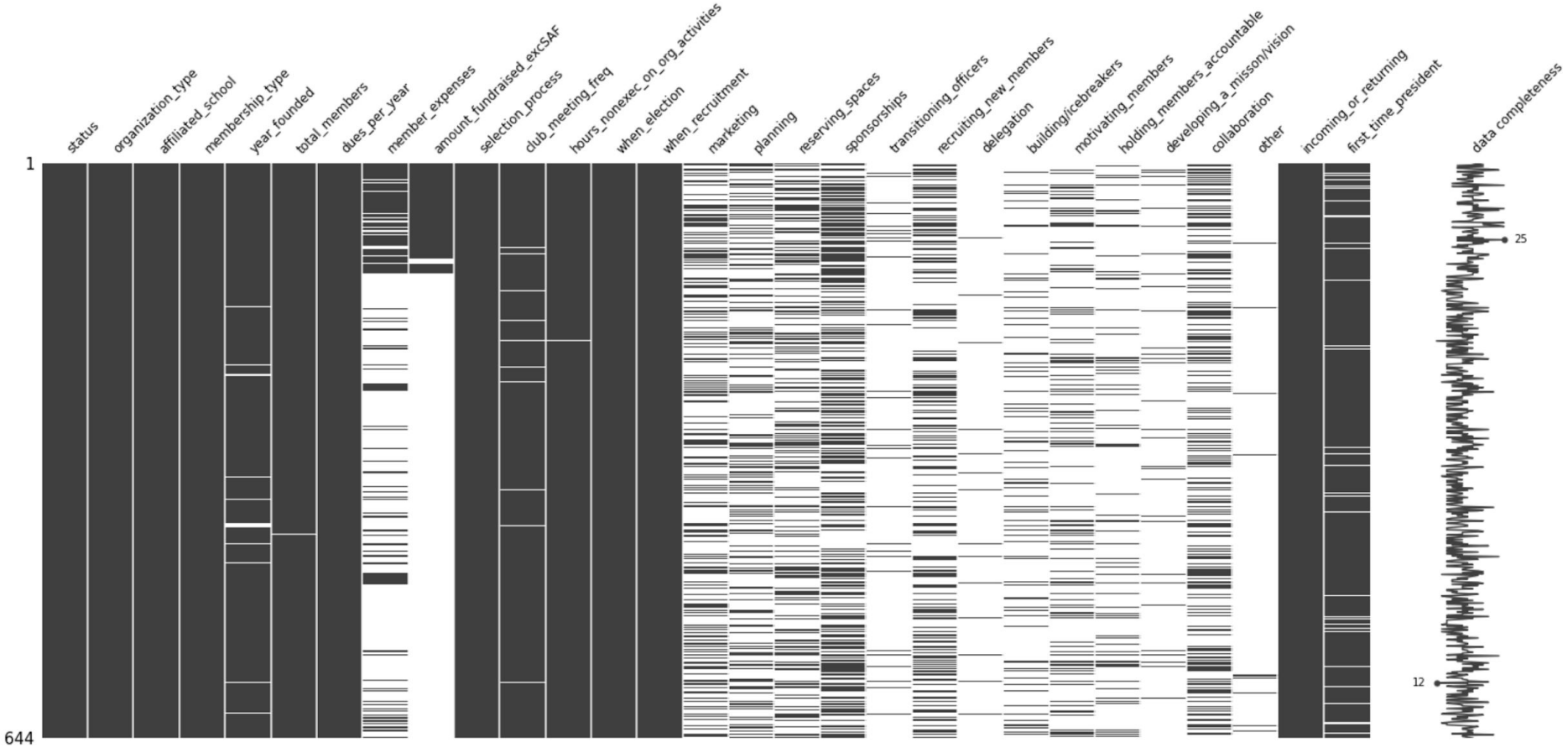
Cleaning

The data was not in a form where we could immediately do analysis

Old Data

	DateCreated	Status	OrganizationType	Name	ShortName	Affiliated School	Membership Type	Approximate year the organization was founded at UVA	Please list 3-5 Primary Events/Programs that your CIO hosts each year. If you do not have annual recurring events,please list some of the events/programs you hosted this past year.	Total number of Members (approximate)	Approximat membershi dues pe academi year (i zero,pleas writ "zero"
0	4/20/20 17:37	Approved	Contracted Independent Organization	Virginia Homecomings	Homecomings	No affiliated schools	Predominantly Undergraduate Students	1995	#ERROR!	45	ze
1	4/23/20 21:42	Approved	Contracted Independent Organization	Sales and Trading Group at McIntire	STGM	Commerce (McIntire)	Predominantly Undergraduate Students	2010	Presentations from faculty, professionals and ...	60	Ze
2	4/24/20 13:52	Approved	Contracted Independent Organization	SpongeBob Club	NaN	No affiliated schools	Predominantly Undergraduate Students	2016	We have a recurring meeting every other week. ...	50	ze
3	4/29/20 21:14	Approved	Contracted Independent Organization	Omega Chi Epsilon	OXE	Engineering	Predominantly Undergraduate Students	2005	1. Tutoring and other academic services2. Prof...	25	3
4	4/18/20 14:33	Approved	Contracted Independent Organization	Alpha Chi Sigma Co-ed Professional Chemistry F...	AXE	Arts and Sciences	Predominantly Undergraduate Students	1922	Hosting speakers for the chemistry department;...	30	$20

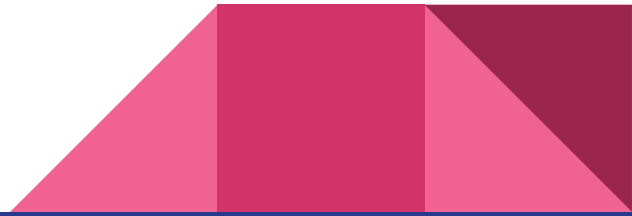
<matplotlib.axes._subplots.AxesSubplot at 0x7f4e43c81d68>



Cleaning

Our aim was to do K-Means on the CIOs since a clear pattern could not be immediately identified in the dataset (if there was one).

- We renamed the columns to shorter names
- We dropped unnecessary columns like DateCreated and Name
- We used try-except routines to change the data that could easily be corrected with code, then exported the edited data frame to a csv and cleaned it more there
- Read the csv back in and prepared it for K-Means



Code Examples for Cleaning

```
#Changing the additional resources columns to dummies
use_ml["marketing"] = pd.get_dummies(use_ml["marketing"])
use_ml["planning"] = pd.get_dummies(use_ml["planning"])
use_ml["reserving_spaces"] = pd.get_dummies(use_ml["reserving_spaces"])
use_ml["sponsorships"] = pd.get_dummies(use_ml["sponsorships"])
use_ml["transitioning_officers"] = pd.get_dummies(use_ml["transitioning_officers"])
use_ml["recruiting_new_members"] = pd.get_dummies(use_ml["recruiting_new_members"])
use_ml["delegation"] = pd.get_dummies(use_ml["delegation"])
use_ml["building/icebreakers"] = pd.get_dummies(use_ml["building/icebreakers"])
use_ml["motivating_members"] = pd.get_dummies(use_ml["motivating_members"])
use_ml["holding_members_accountable"] = pd.get_dummies(use_ml["holding_members_accountable"])
use_ml["developing_a_misson/vision"] = pd.get_dummies(use_ml["developing_a_misson/vision"])
use_ml["collaboration"] = pd.get_dummies(use_ml["collaboration"])
use_ml["other"] = pd.get_dummies(use_ml["other"])
```

```
for i in use_ml["hours_nonexec_on_org_activities"]:
  try:
    use_ml["hours_nonexec_on_org_activities"].replace(i, int(i), inplace=True)
  except:
    try:
      updated = ""
      if i[2:5] in months:
        lower = months[i[2:5]]
        higher = i.strip(i[1:5])
        use_ml["hours_nonexec_on_org_activities"].replace(i, (int(lower) + int(higher)) / 2, inplace=True)
      elif i[3:6] in months:
        lower = months[i[3:6]]
        higher = i.strip(i[2:6])
        use_ml["hours_nonexec_on_org_activities"].replace(i, (int(lower) + int(higher)) / 2, inplace=True)
      else:
        try:
          if i.lower() in change:
            use_ml["hours_nonexec_on_org_activities"].replace(i, 0, inplace=True)
          else:
            print(i)
        except:
          if str(i).lower() == "nan":
            use_ml["hours_nonexec_on_org_activities"].replace(i, 0, inplace=True)
          else:
            print(i)
    except:
      print(i)
```

Cleaning

The cleaned dataset

	organization_type	affiliated_school	membership_type	total_members	dues_per_year	member_expenses	amount_fundraised_excSAF	selection_process	club_meeting_freq	hours_n
0	Contracted Independent Organization	No affiliated schools	Predominantly Undergraduate Students	45.0	0	0	0	Neither	Weekly	
1	Contracted Independent Organization	Commerce (McIntire)	Predominantly Undergraduate Students	60.0	0	0	0	Neither	Weekly	
2	Contracted Independent Organization	No affiliated schools	Predominantly Undergraduate Students	50.0	0	0	0	Neither	Weekly	
3	Contracted Independent Organization	Engineering	Predominantly Undergraduate Students	25.0	30	0	0	Neither	Once or twice a semester	
4	Contracted Independent Organization	Arts and Sciences	Predominantly Undergraduate Students	30.0	200	0	0	Application	Weekly	
5	Contracted Independent Organization	No affiliated schools	Both Undergraduate and Graduate Students	80.0	6100	0	0	Neither	Weekly	
6	CIO - Law	Law	Predominantly Graduate Students	200.0	10	0	0	Neither	Monthly	

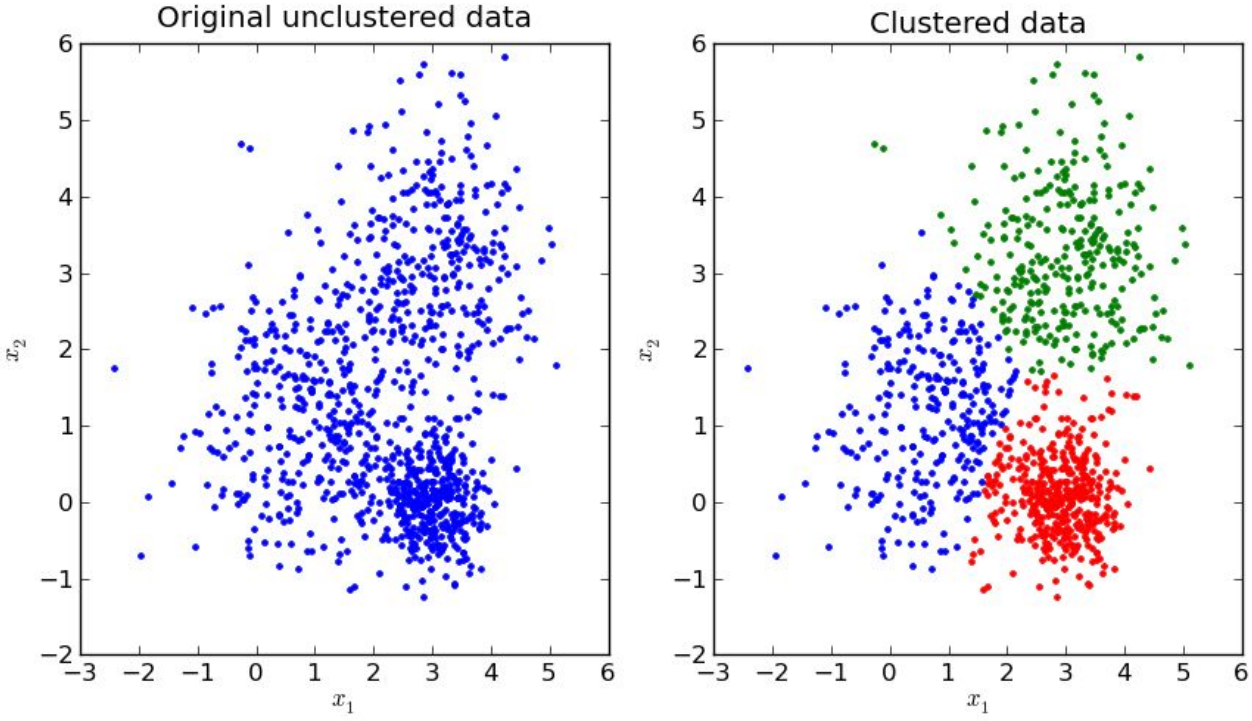
Clustering

We aimed to find any kind of similarity between the CIOs by using K-Means to group the CIOs into four clusters.

The columns we ended up using for K-Means were

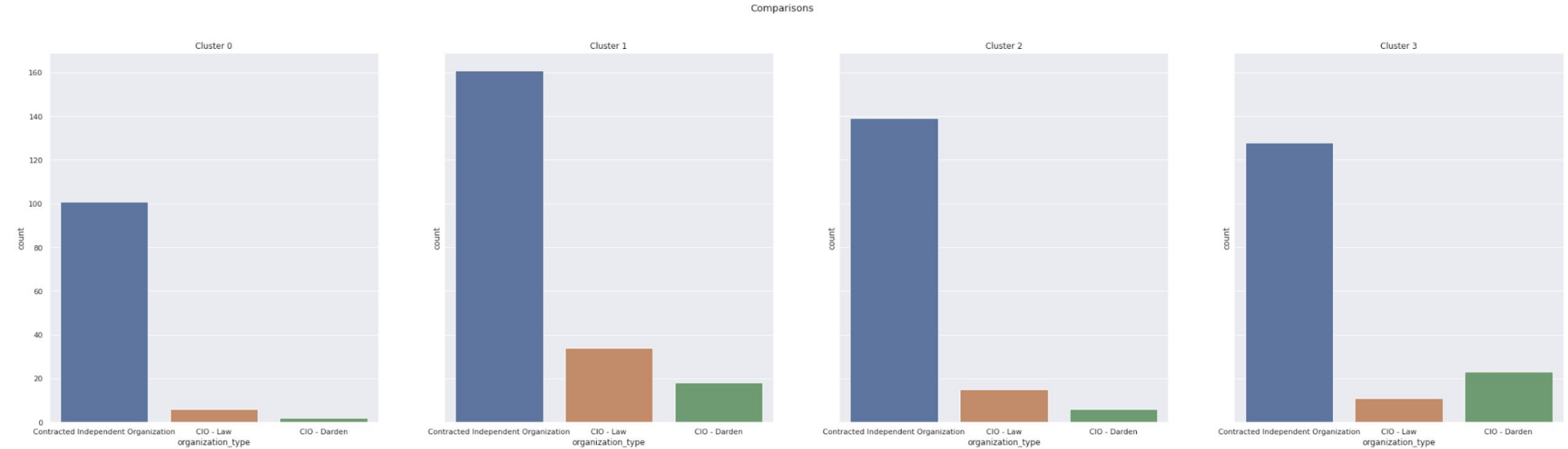
```
total_names = ["name", "organization_type", "affiliated_school", "membership_type", "total_members", "dues_per_year",
    "member_expenses",  "amount_fundraised_excSAF", "selection_process", "club_meeting_freq", "hours_nonexec_on_org_activities",
    "when_election", "when_recruitment",  "marketing", "planning",  "reserving_spaces", "sponsorships", "transitioning_officers",
    "recruiting_new_members", "delegation", "building/icebreakers", "motivating_members", "holding_members_accountable",
    "developing_a_misson/vision", "collaboration", "other", "incoming_or_returning", "first_time_president"]
```

What is K-Means

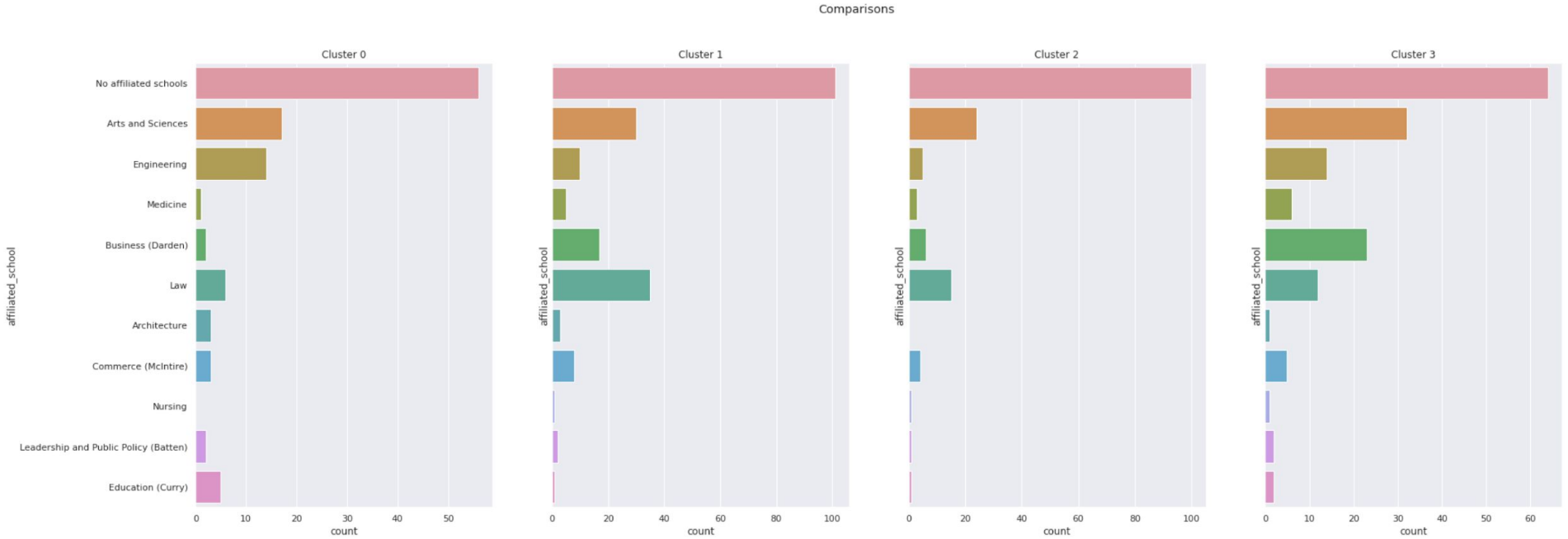


Visualizations

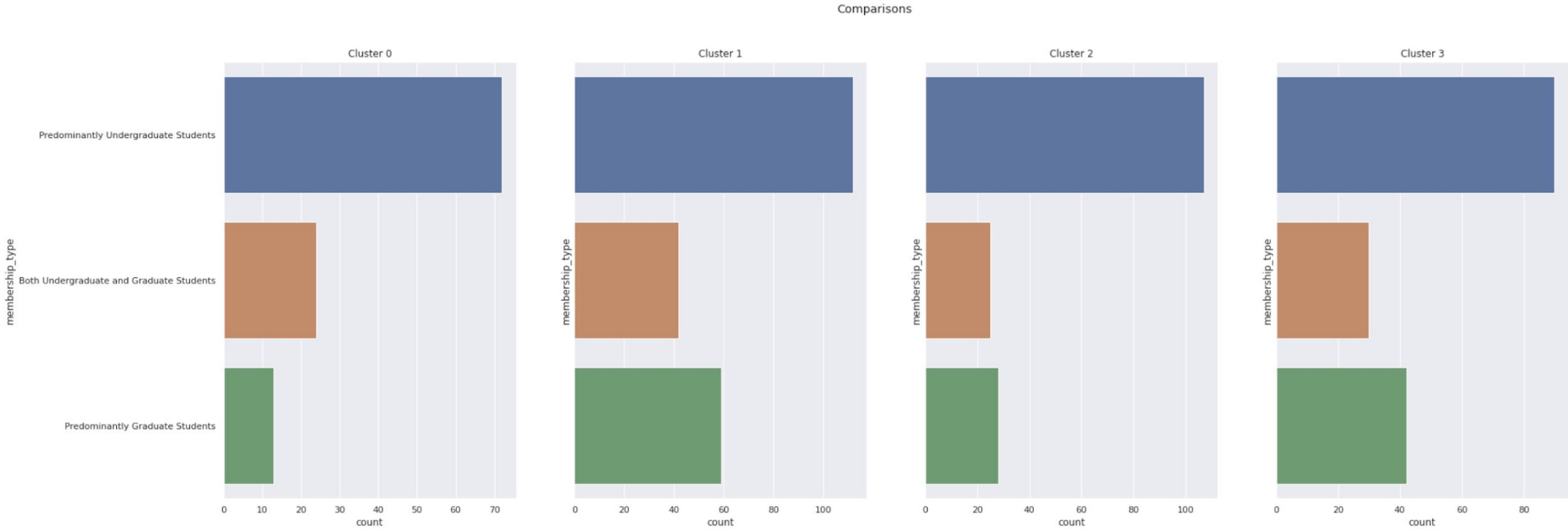
Organization Type



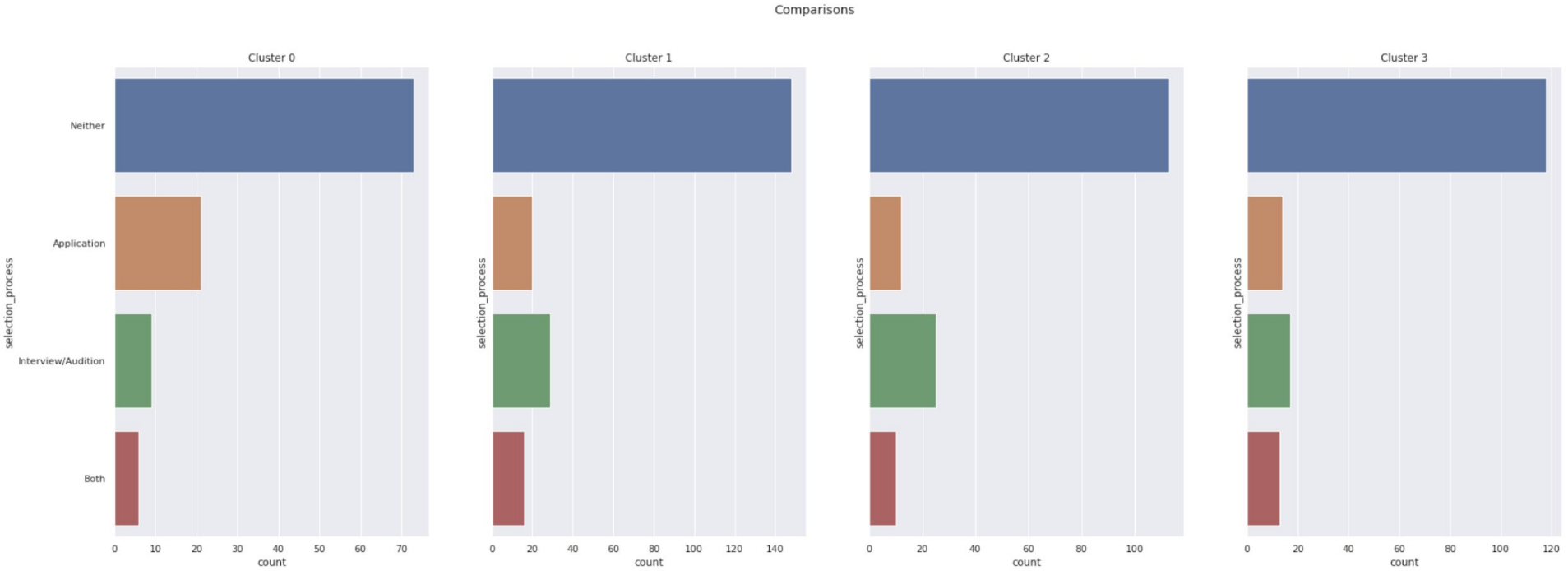
Affiliated School



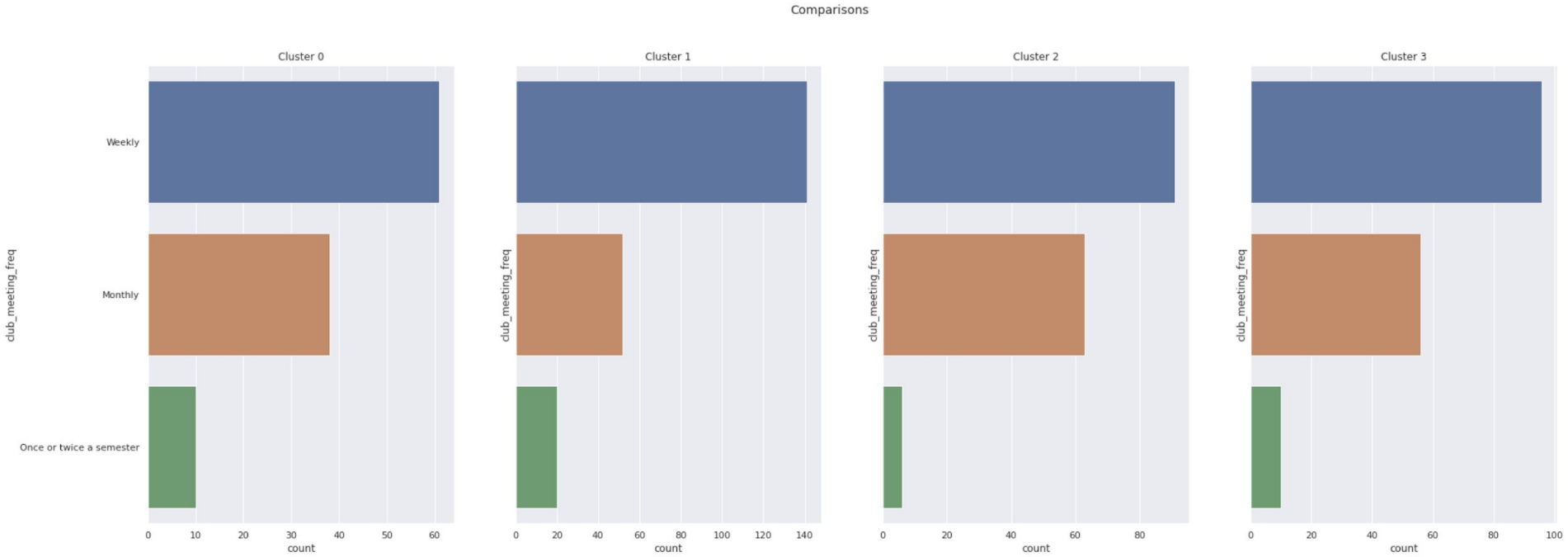
Membership Type



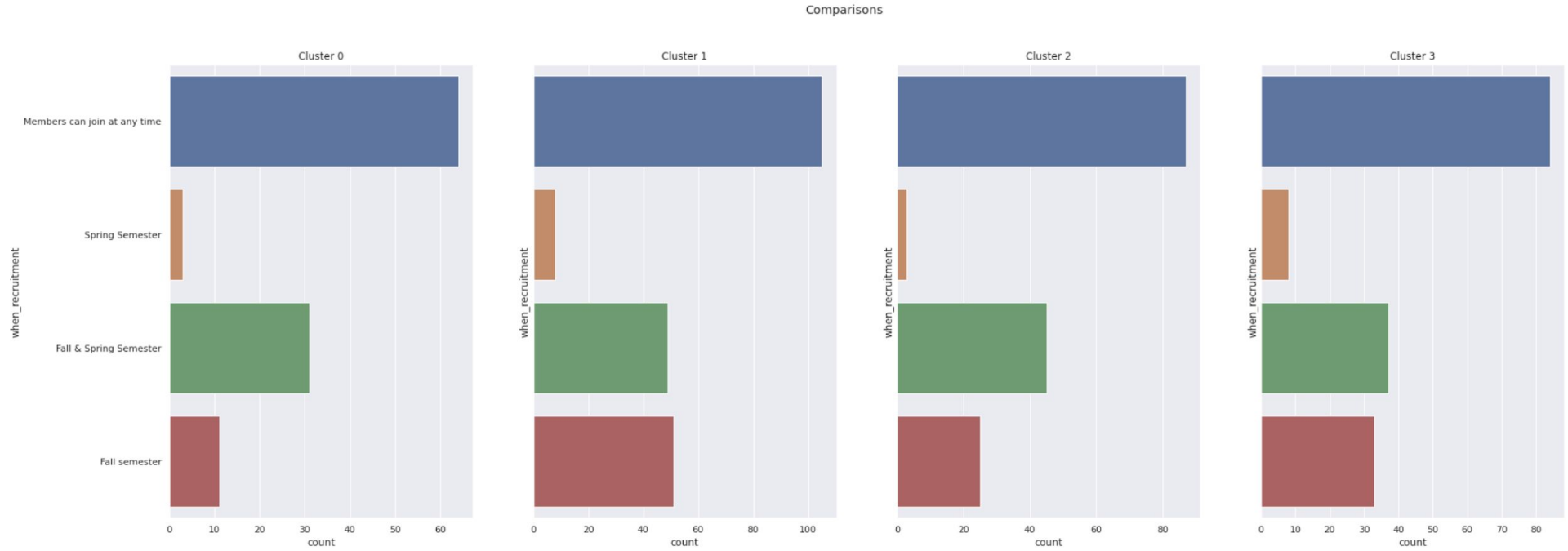
Selection Process



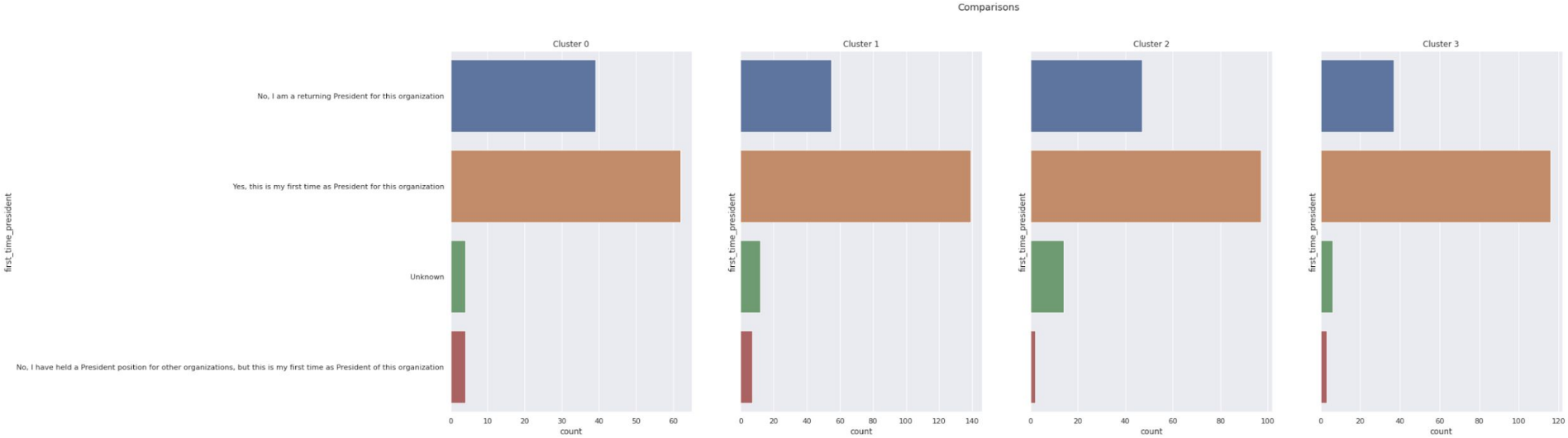
Club Meeting Frequency



Recruitment start date

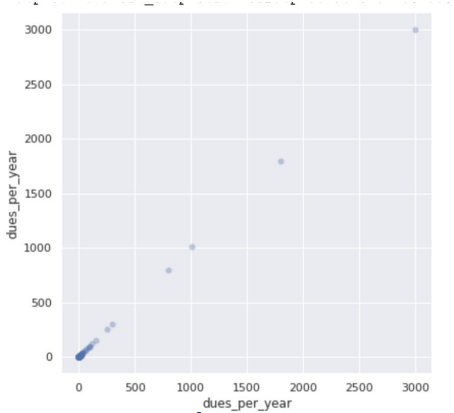


First Time President?



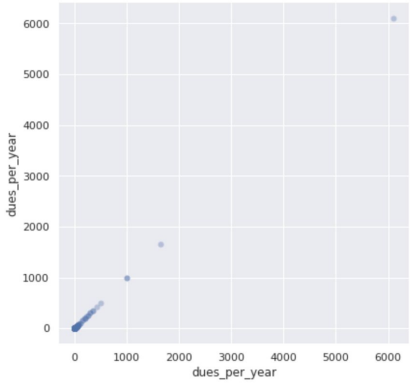
Dues Per Year

```
count     109.000000
mean       78.376147
std       353.196273
min         0.000000
25%         0.000000
50%         0.000000
75%        20.000000
max      3000.000000
Name: dues_per_year, dtype: float64
```



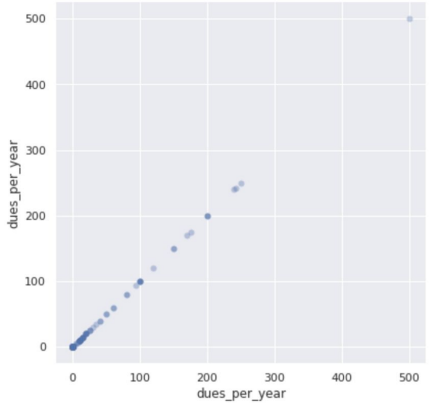
Cluster 0

```
count     213.000000
mean       73.154930
std       445.479347
min         0.000000
25%         0.000000
50%         0.000000
75%        20.000000
max      6100.000000
```



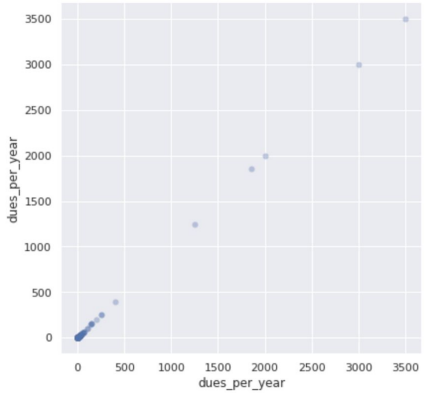
Cluster 1

```
count     160.000000
mean       26.187500
std        64.162254
min         0.000000
25%         0.000000
50%         0.000000
75%        15.000000
max       500.000000
```



Cluster 2

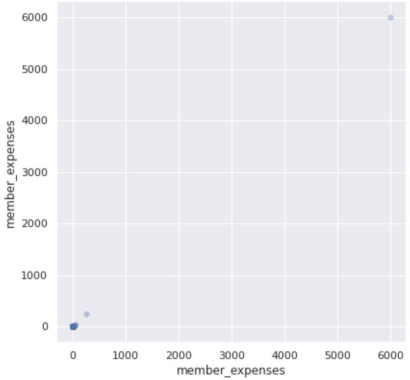
```
count     162.000000
mean       91.018519
std       427.367043
min         0.000000
25%         0.000000
50%         0.000000
75%        20.000000
max      3500.000000
```



Cluster 3

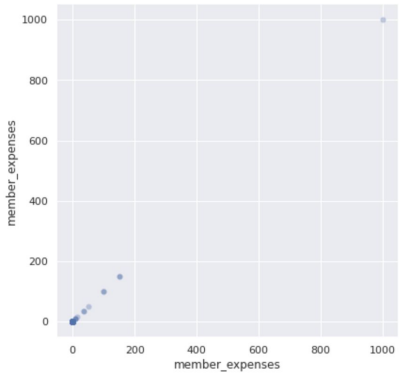
Member Expenses

count	109.000000
mean	58.715596
std	574.869310
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	6000.000000



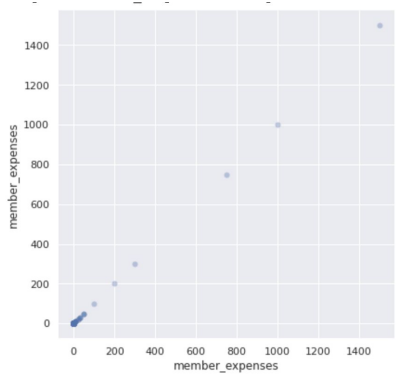
Cluster 0

count	213.000000
mean	7.816901
std	70.625903
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	1000.000000



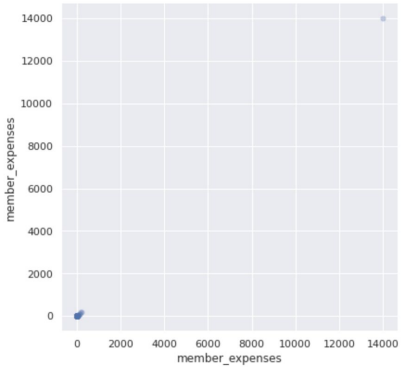
Cluster 1

count	160.000000
mean	26.406250
std	155.672786
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	1500.000000



Cluster 2

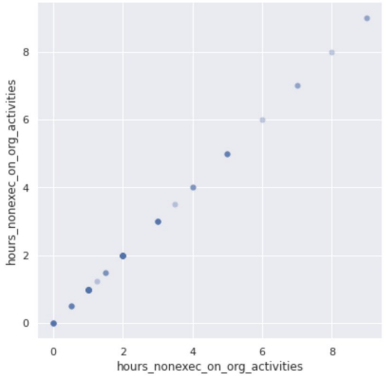
count	162.000000
mean	91.567901
std	1099.798422
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	14000.000000



Cluster 3

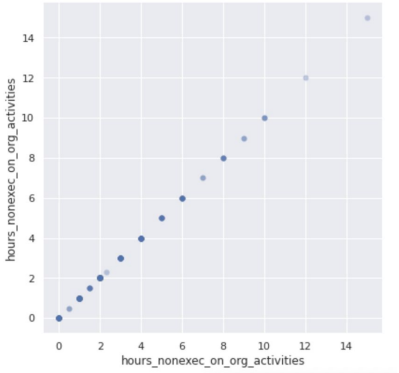
Hours on Non-Executive Organization Activities

count	109.000000
mean	1.942661
std	1.821818
min	0.000000
25%	1.000000
50%	1.000000
75%	2.000000
max	9.000000



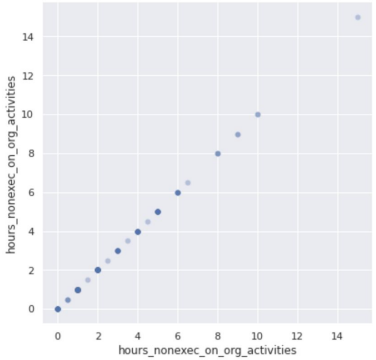
Cluster 0

count	213.000000
mean	2.541396
std	2.405446
min	0.000000
25%	1.000000
50%	2.000000
75%	3.000000
max	15.000000



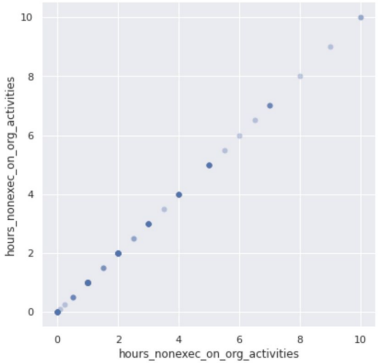
Cluster 1

count	160.000000
mean	2.481250
std	2.357837
min	0.000000
25%	1.000000
50%	2.000000
75%	4.000000
max	15.000000



Cluster 2

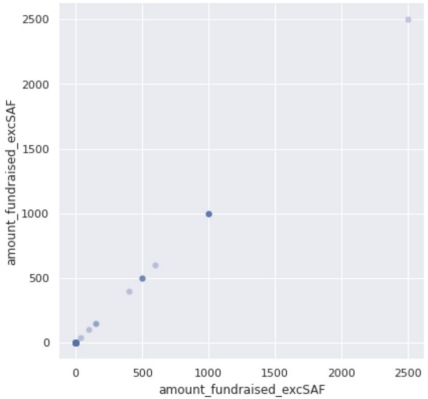
count	162.000000
mean	2.113272
std	1.997816
min	0.000000
25%	1.000000
50%	1.500000
75%	3.000000
max	10.000000



Cluster 3

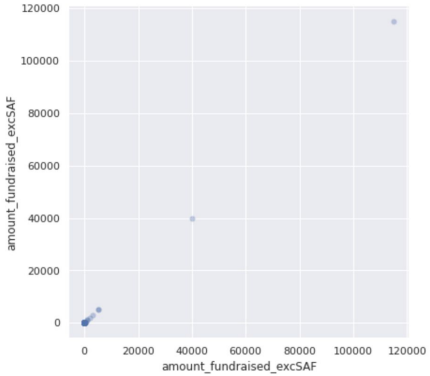
Amount Fundraised Except from SAF

count	109.000000
mean	109.541284
std	340.460977
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	2500.000000



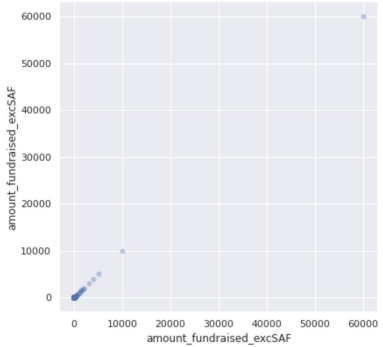
Cluster 0

count	213.000000
mean	814.464789
std	8341.080536
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	115000.000000



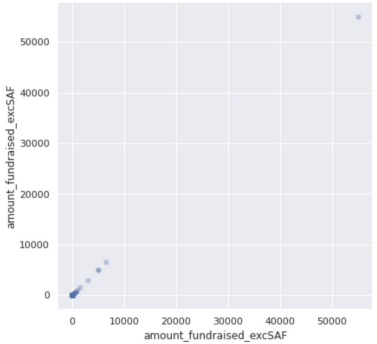
Cluster 1

count	160.000000
mean	615.800000
std	4830.087276
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	60000.000000



Cluster 2

count	162.000000
mean	500.987654
std	4381.502934
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	55000.000000



Cluster 3

Average Total Membership in Clubs for Each Cluster

```
Total Members in Clubs: 50363

 Average Total Members Cluster 0:  75.46976777522936
 Average Total Members Cluster 1:  93.08173415492958
 Average Total Members Cluster 2:  50.757529296875
 Average Total Members Cluster 3:  87.5925925925926
```

Insights

How many clubs are students involved with on average?

Student Population: Fall 2020

Undergraduate	Graduate	Total
17,314	8,328	25,642

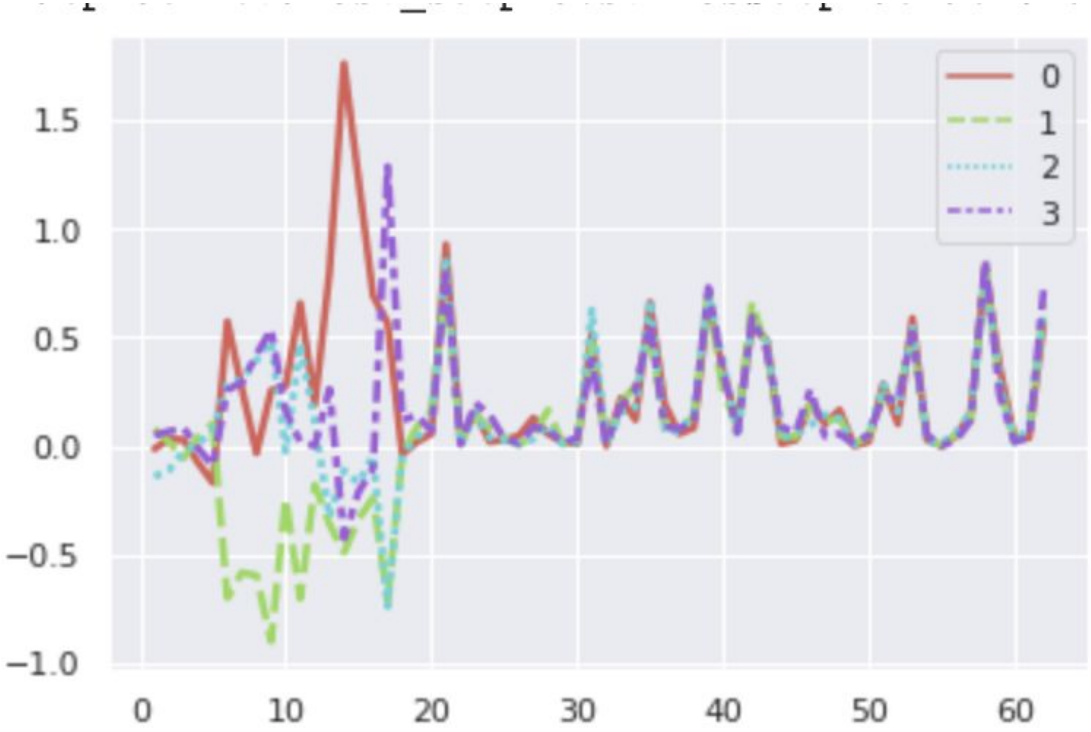
	#Clubs
Total Predominantly Graduate Student Clubs Membership: **13,106**	**142**
Total Predominantly Undergraduate Student Clubs Membership: **28,317**	**381**
Total Both Undergrad and Grad Student Clubs Membership: **8,940**	**121**

Grad Student Clubs = 13106 / 8328 = **1.57 clubs per grad student**
Undergrad Student Clubs = 28317 / 17314 = **1.64 clubs per undergrad student**

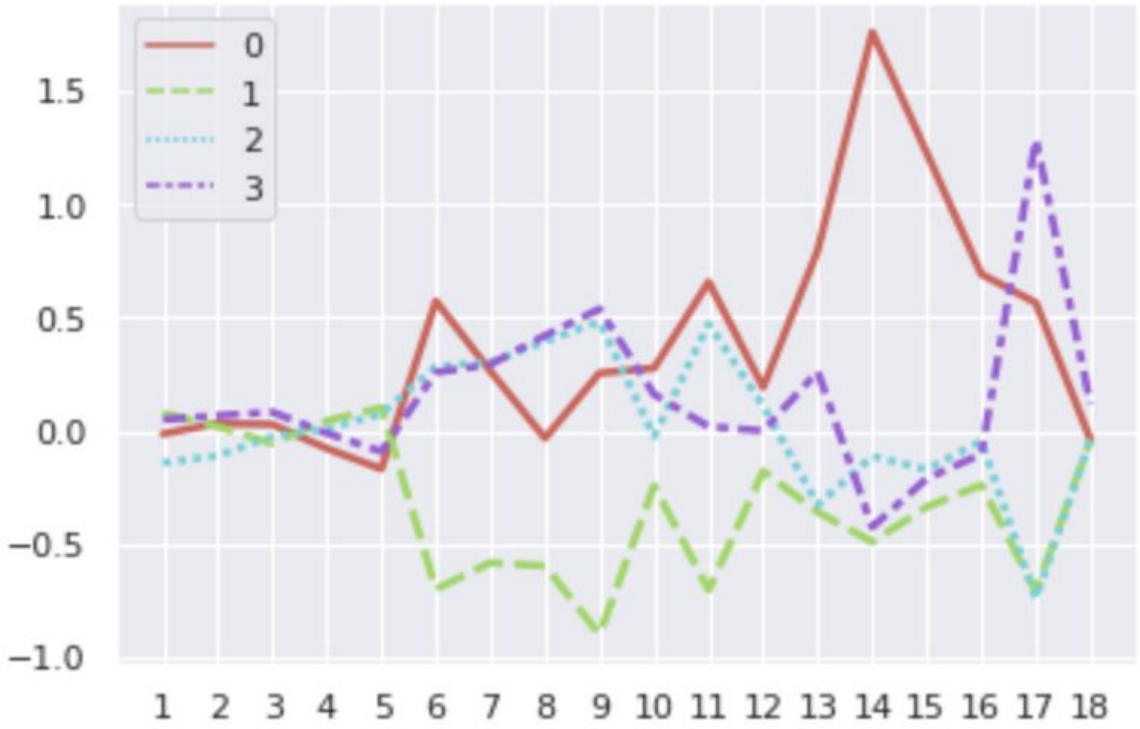
Assuming ⅔ undergrad, ⅓ grad makeup:
Grad Student Clubs = (13106 + (⅓ * 8940) / 8328 = **1.93 clubs per grad student**
Undergrad Student Clubs = (13106 + (⅔ * 8940) / 8328 = **1.98 clubs per undergrad student**

Overall Participation: **1.96 clubs per student**

Each Cluster's Normalized Mean For Each Feature



Organization Type: 19-21
Affiliated School: 22-32
Membership Type: 33-35
Selection Process: 36-39
Club Meeting Frequency: 40-42
When's the Election: 43-52
When's Recruiting: 53-56
Incoming or Returning Pres: 57-59
First-Time President: 60-62




```
{1: 'total_members',
 2: 'dues_per_year',
 3: 'member_expenses',
 4: 'amount_fundraised_excSAF',
 5: 'hours_nonexec_on_org_activities',
 6: 'marketing',
 7: 'planning',
 8: 'reserving_spaces',
 9: 'sponsorships',
 10: 'transitioning_officers',
 11: 'recruiting_new_members',
 12: 'delegation',
 13: 'building/icebreakers',
 14: 'motivating_members',
 15: 'holding_members_accountable',
 16: 'developing_a_misson/vision',
 17: 'collaboration',
 18: 'other'}
```

name
The Machine Learning Club at the University of...
qMD
Arts Grounds Guides
Greens to Grounds
The Student Chapter of the American Chemical S...
The Roosevelt Society
Omicron Delta Kappa
Photography Club at UVA
My Sister's Keeper
Doctors Without Borders at UVA
University Peer Advising Link
International Relations Organization at the Un...
Chapter of the National Organization of Minori...
Catholic Hoos at the University of Virginia
Mind&Body at UVA
Virginia Journal of Law & Technology
Rise Together
The Monroe Society at UVA
Parkour Association
LabShorts
Omega Chi Epsilon
Washington Literary Society and Debating Union
Virginia Law and Business Society at The Unive...
Health Law Association at UVA
College Diabetes Network at the University of ...
undocUVA
Virginia Case Club
Virginia Law Democrats
University Democrats at UVA
The Lebanese Club at the University of Virgini...
Outdoors Club at Darden
Consulting Club at Darden
The University Fellowship at St. Paul's
Law Christian Fellowship
Amuse Bouche
Dream Corps International
League of Legends at UVA
The Wesley Foundation at UVA
Unsung People
Society for the Promotion of Indian Classical ...
Virginia Homecomings
Sales and Trading Group at McIntire
Afro-Latinx Student Organization
Student Organization for Caribbean Awareness
180 Degrees Consulting at UVA
Men's Water Polo Club at the University of Vir...
Student Virginia Education Association
Circle K International
Japan Club
College Science Scholars Council
Lutheran Campus Ministry
Pulsar Observation Group at UVA
Shakespeare On the Lawn
The Fralin Student Docents at the University o...
Fill The Hill
Latter-day Saint Student Association
Virginia Interfaith Coalition
Mock Trial at UVA Law
Smash At UVa
Chinese Christian Student Fellowship at UVA
SpongeBob Club
Alpha Chi Sigma Co-ed Professional Chemistry F...
Gizmologists
Spikeball Club at the University of Virginia (...
Shotokan Karate Club at The University of Virg...
One for the World at UVA
Pakistani Students' Association
Changing Health Attitudes & Actions to Recreat...
Aerial Dance Club at UVA
National Association for the Advancement of Co...
Astronomy Club at the University of Virginia
Historical Simulation Society
Latinx Leadership Coalition
INCOSE at the University of Virginia
Collegiate 100 Society of Virginia
Young Americans for Freedom at UVA
Virginia South Asian Law Students Association
Muslim Students Association
National Society of Black Engineers at UVA
Virginia Motorsports

Cluster 0

name
The Machine Learning Club at the University of...
qMD
Arts Grounds Guides
Greens to Grounds
The Student Chapter of the American Chemical S...
The Roosevelt Society
Omicron Delta Kappa
Photography Club at UVA
My Sister's Keeper
Doctors Without Borders at UVA
University Peer Advising Link
International Relations Organization at the Un...
Chapter of the National Organization of Minori...
Catholic Hoos at the University of Virginia
Mind&Body at UVA
Virginia Journal of Law & Technology
Rise Together
The Monroe Society at UVA
Parkour Association
LabShorts

Highest Features

- Motivating members
- Holding members accountable
- Developing mission/vision
- Recruiting new members
- Marketing

Cluster 1

Highest Feature:

- Either well-off or inactive

Cluster 2

Virginia Homecomings	Lutheran Campus Ministry
Sales and Trading Group at McIntire	Pulsar Observation Group at UVA
Afro-Latinx Student Organization	Shakespeare On the Lawn
Student Organization for Caribbean Awareness	The Fralin Student Docents at the University o...
180 Degrees Consulting at UVA	Fill The Hill
Men's Water Polo Club at the University of Vir...	Latter-day Saint Student Association
Student Virginia Education Association	Virginia Interfaith Coalition
Circle K International	Mock Trial at UVA Law
Japan Club	Smash At UVa
College Science Scholars Council	Chinese Christian Student Fellowship at UVA

Highest Feature:

- Recruiting Members

Cluster 3

Highest Features:

- Collaboration
- building/icebreakers
- Sponsorships

Cluster 0:

- Motivating members
- Holding members accountable
- Developing mission/vision
- Recruiting new members
- Marketing

Cluster 1:

- Either well-off or inactive

Cluster 2:

- Recruiting Members

Cluster 3:

- Collaboration
- building/icebreakers
- Sponsorships

Clubs in Cluster 0 seem to want to grow and maintain members.

Clubs in Cluster 2 seem established, but want to grow.

Clubs in Cluster 3 seem to want additional resources on team-building activities. These clubs could be looking to engage members of their club. This could also mean that these clubs want to place more emphasis on inclusion.

Thank You!