Features of Startup Founders

— in the US Health Care Industry —

4/13/2017 Mei Fu

QUESTION & MOTIVATION

Startups are important

+ they take a leading role in the innovation

Founders are important

+ the success of a startup is largely related to them

+ relatively hard to quantitatively analysis their features through traditional databases

This project is interested in illustrating some brief common features of founders who have a higher probability of success or more favored by venture capitals (raised at least one round)

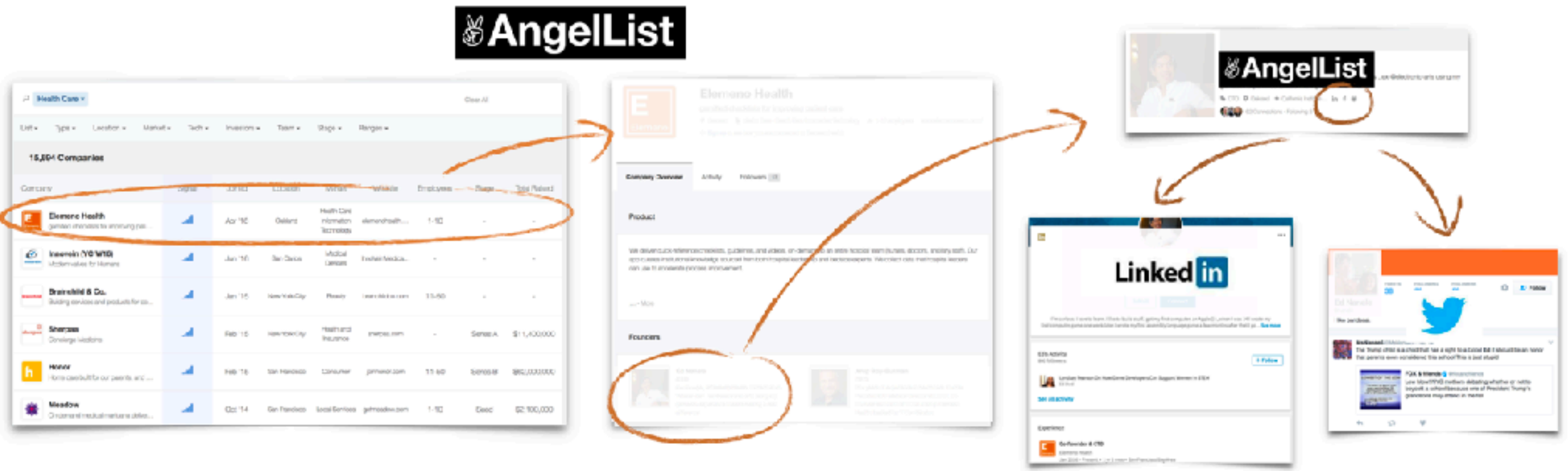
DATASETS - ANGELLIST, LINKEDIN, TWITTER

+ **Networks**: former companies and schools
+ **Skill-sets**: former position, major, and self-reported skills
+ **Influence**: tweets, followers, retweets, likes

Step1: Startup Filtering
Basic Information

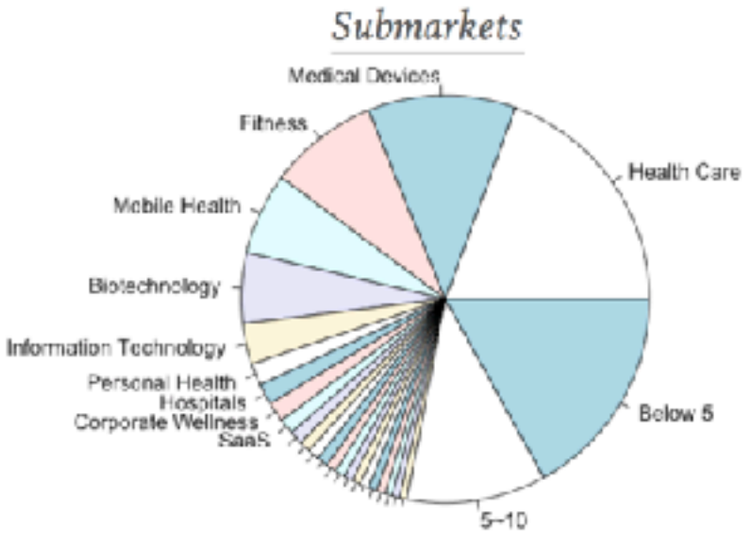
Step2: Startup Info
Funding Details

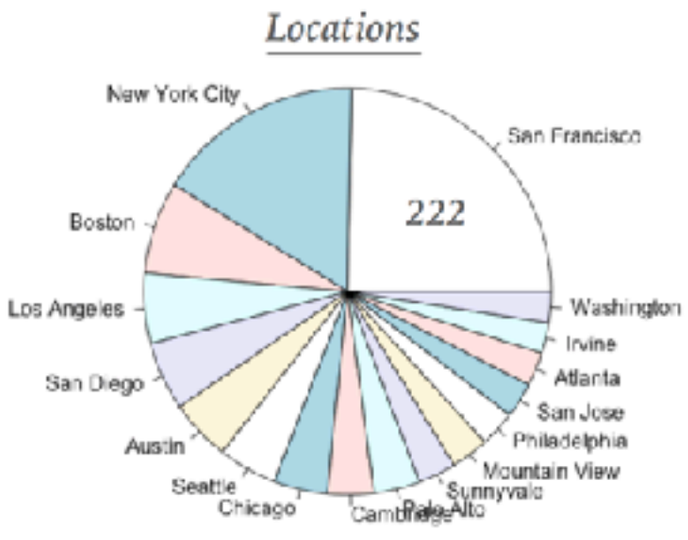
Step3: Founder Info
Networks, Skills, Influence

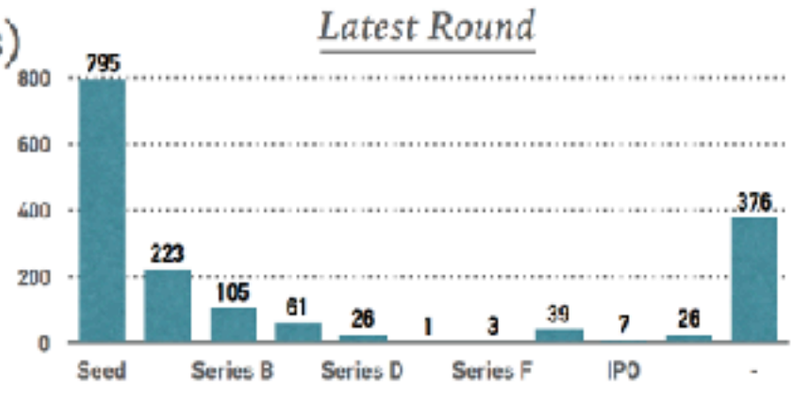


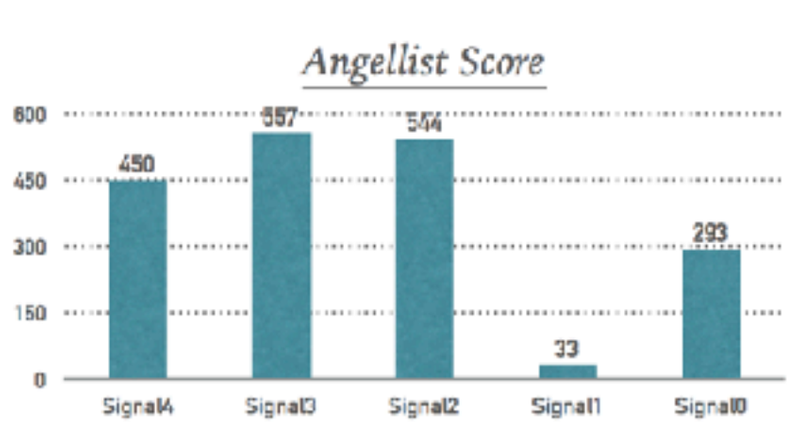
DATASETS - FILTERING CRITERIA

- **Industry**: Health Care (15,693 Companies)
- **Filter Range**: US + Having Raising Amount (1,929 Companies)









- **Selected Range**: (200+)

4 QUESTIONS

Q1: How’s their social networks?

+ Do they form communities or have close relationships

Q2: What are the most frequent schools or companies?

Q3: What are they good at?

+ What are the most important / popular skills or experiences to be a good founder?

Q4: Social media vs funding rounds?

Q1: HOW'S THEIR SOCIAL NETWORKS?

Data Processing Steps

- Load & Clean Data; CSV —> Dataframe

1

```
ad data
ool.data <- data.table(read.csv('startups_founder_networks_total_schools.

# clean data
school.cleaned <- list()
count <- 0
for (i in 1:nrow(school.data)) {
  startup <- as.character(school.data[i,1][[1]])
  founder1 <- as.character(school.data[i,2][[1]])
  schools <- str_split(school.data[i,3][[1]], '\\|')
  for (j in 1:length(schools[[1]])) {
    count <- count + 1
    school.cleaned[[count]] = c(startup, founder1, schools[[1]][j])
  }
}
school.cleaned <- as.data.frame(school.cleaned)
school.cleaned <- as.data.frame(t(school.cleaned))
names(school.cleaned) <- c('Startup', 'Founder.Name', 'School')
```

- Build 2 node sets — founders & schools
- Combine them by their relationship

2

```
ild founders nodeset on one side
under_set1 <- as.data.frame(unique(school.cleaned[,2]))
founder_set1 <- cbind(founder_set1, c(1:nrow(founder_set1)))
names(founder_set1) <- c('Founder.Name','Founder.ID')
founder_nodeset1 <- founder_set1$Founder.ID

# build schools nodeset on the other side
school_set <- as.data.frame(unique(school.cleaned[,3]))
school_set <- cbind(school_set, c(1:nrow(school_set)))
names(school_set) <- c('School','School.ID')
school_nodeset <- school_set$School.ID

# build their relationship
school.School <- full_join(school_set, school.cleaned[,2:3], by = 'School')
school.School <- full_join(founder_set1, school.School, by = 'Founder.Name')
school.draw <- data.frame(school.School$School.ID, school.School$Founder.ID)
names(school.draw) <- c('School.ID', 'Founder.ID')
```

- Use igraph to generate the network
- Build a bipartite projection on the founder's side

3

```
e igraph to draw the bipartite graph and its projection graph
g <- graph.empty()
g <- add.vertices(g, nv=length(founder_nodeset1), attr=list(name=paste0('A',
type=rep(TRUE,length(founder_nodeset1)))
g <- add.vertices(g, nv=length(school_nodeset), attr=list(name=paste0('B',sch
type=rep(FALSE,length(school_nodeset)))
edgeListVec <- as.vector(t(as.matrix(data.frame(S1=paste0('A',school.draw$Fou
S2=paste0('B',school.draw$School.ID)))))
g <- add.edges(g, edgeListVec)

proj <- bipartite_projection(g, multiplicity = TRUE)
```

- Use walktrap to illustrate communities
- Use R shiny to draw an interactive graph with different random walk steps for walktrap

4

```
walktrap to illustrate the potential communities
use R shiny to draw an interactive graph with different random walk steps for walktrap
tput$school <- renderPlot({
  m <- cluster_walktrap(proj1[[2]], weights = E(proj1[[2]])$weight, steps = input$number, merges = T
RUE, modularity = TRUE, membership = TRUE)
  plot(proj1[[2]], layout = layout.fruchterman.reingold,
       vertex.size = 3, vertex.color = adjustcolor(membership(m), alpha.f = 0.6),
       vertex.frame.color = adjustcolor('grey', alpha.f = 0.5),
       vertex.label = NA,
       vertex.label.cex = 1.2, vertex.label.font = 2, vertex.label.color = adjustcolor('Black', a
lpha.f = 0.8),          vertex.label.degree = -pi/2,
       edge.arrow.size = 0.8, edge.color = adjustcolor('grey', alpha.f = 0.5))
})

inputPanel(
  sliderInput('number', label = 'Walktrap Steps:', value = 3, min = 1, max = 10)
)

renderUI({
  plotOutput('school', height = 600, width = 600)
})
```

Q1: HOW'S THEIR SOCIAL NETWORKS?

Networks based on Schools...

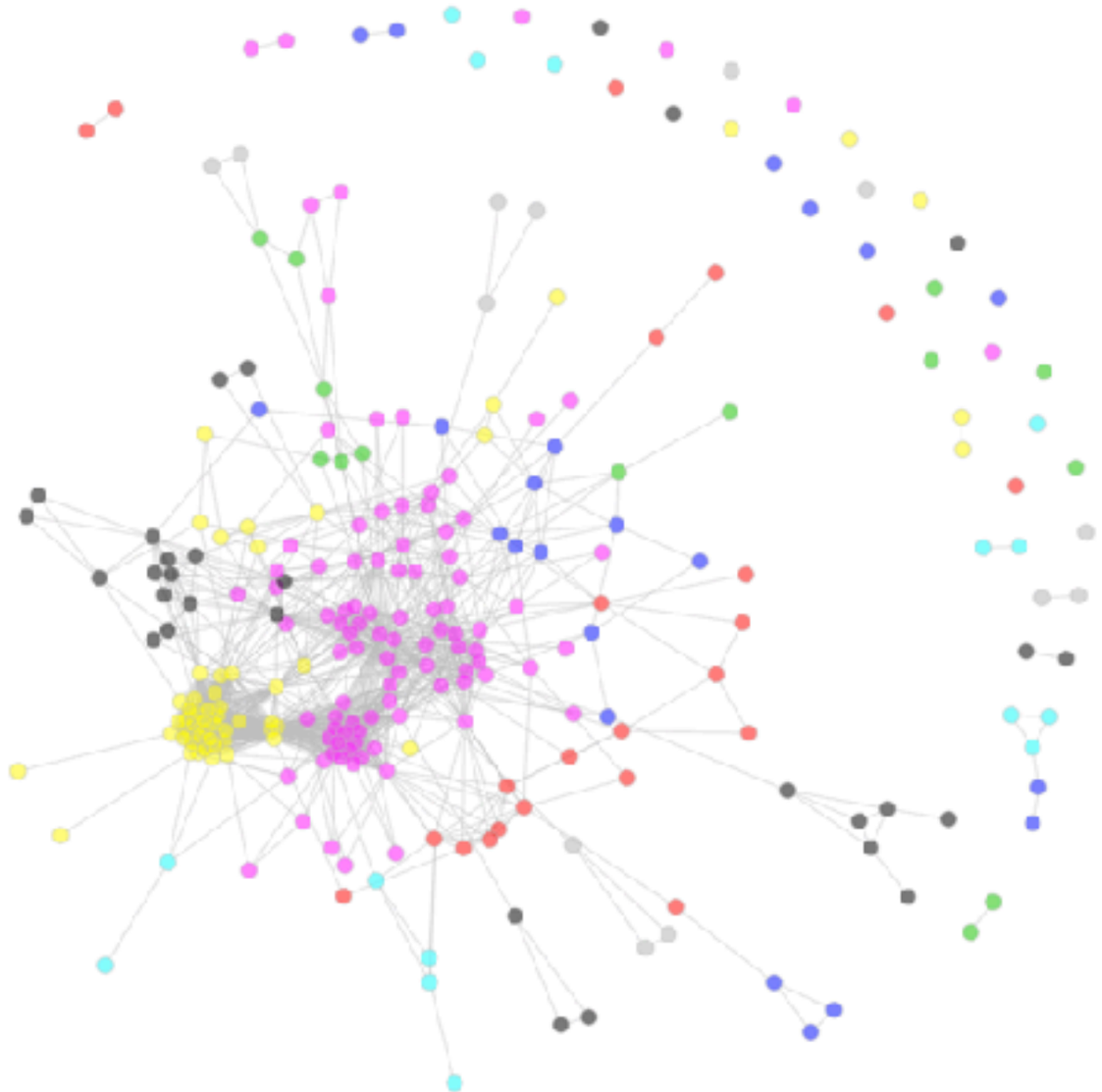
+ 60% founders are connected in the largest group
+ Several sub communities

based on Former Companies...

+ 14% founders are connected in the largest group

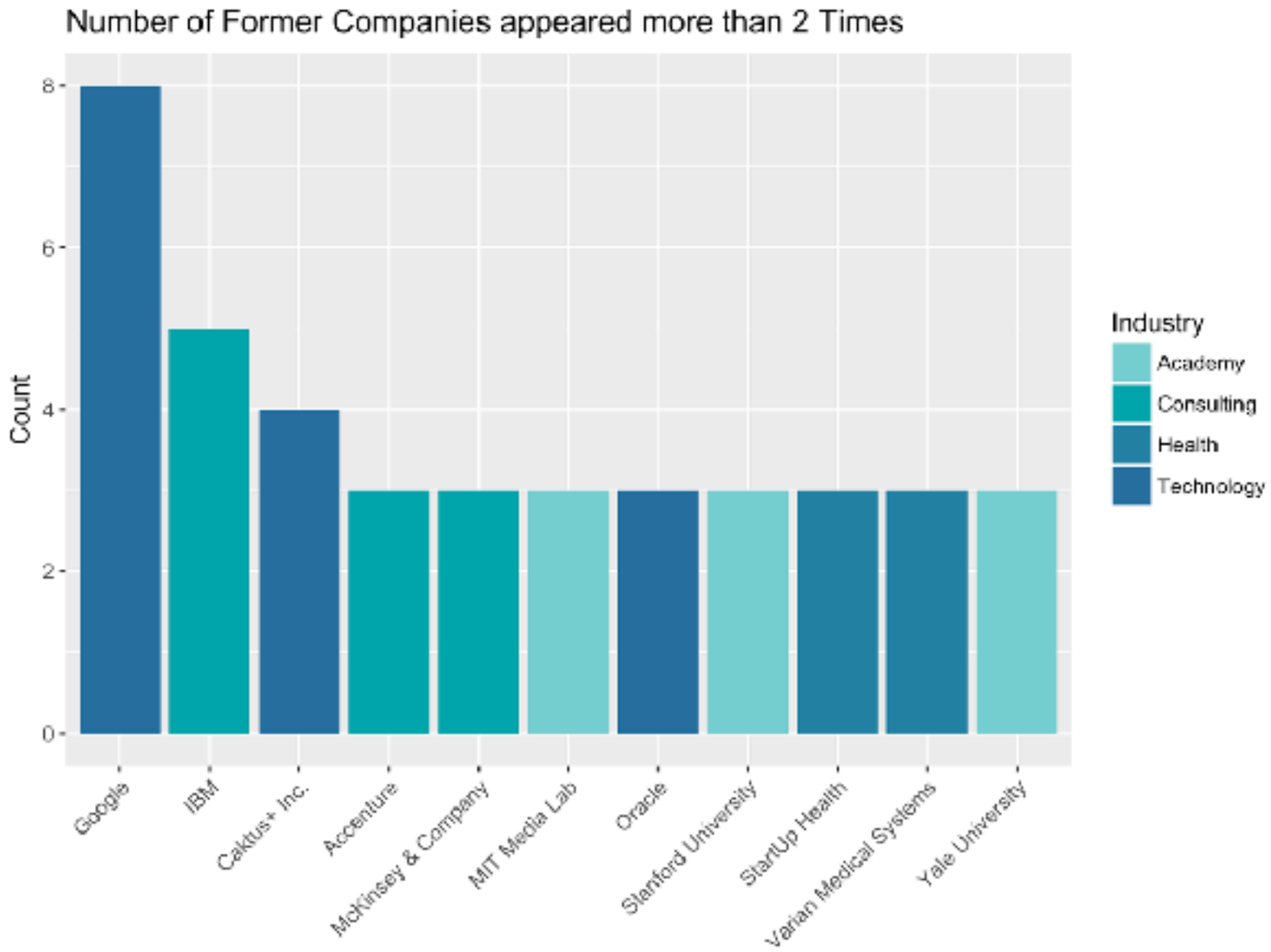
Q1: HOW'S THEIR SOCIAL NETWORKS?

including Schools, Former Companies, and Current startup...



+ 82% founders are connected in the largest group
+ Two major communities
+ Average shortest path is 3.1

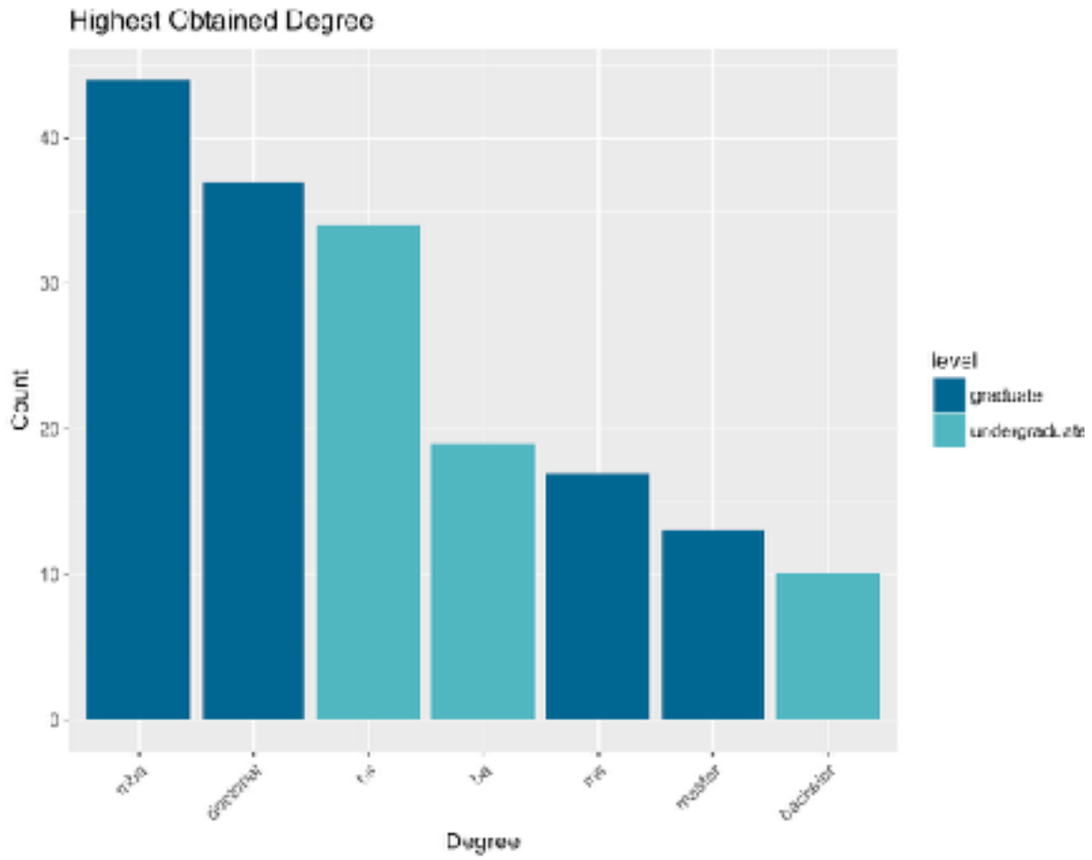
Q2: MOST FREQUENT SCHOOLS AND COMPANIES?



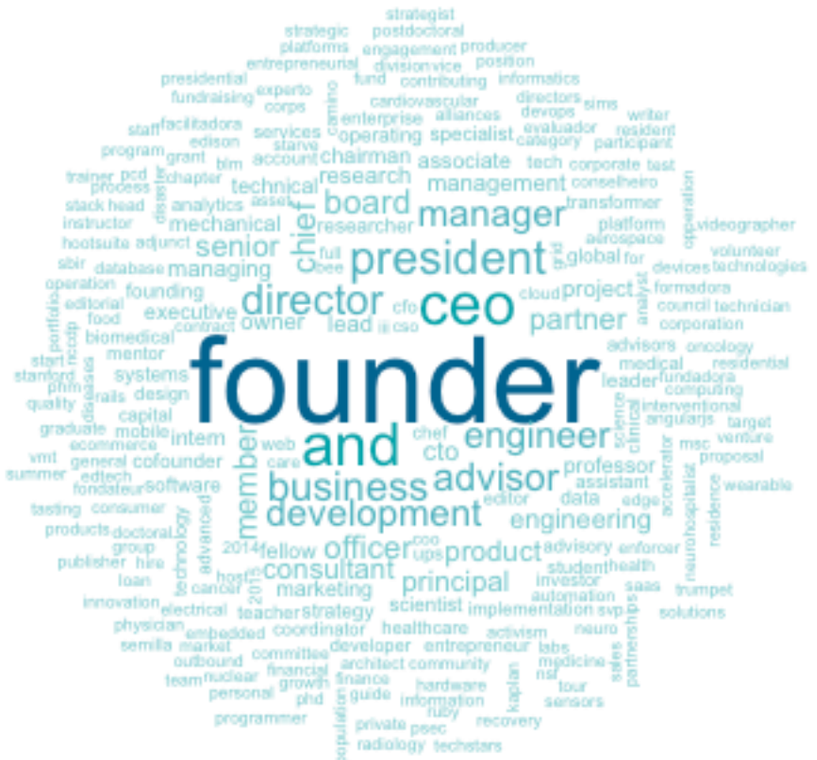
+ Technolog

Q3: WHAT ARE THEY GOOD AT?

64% have graduate degree

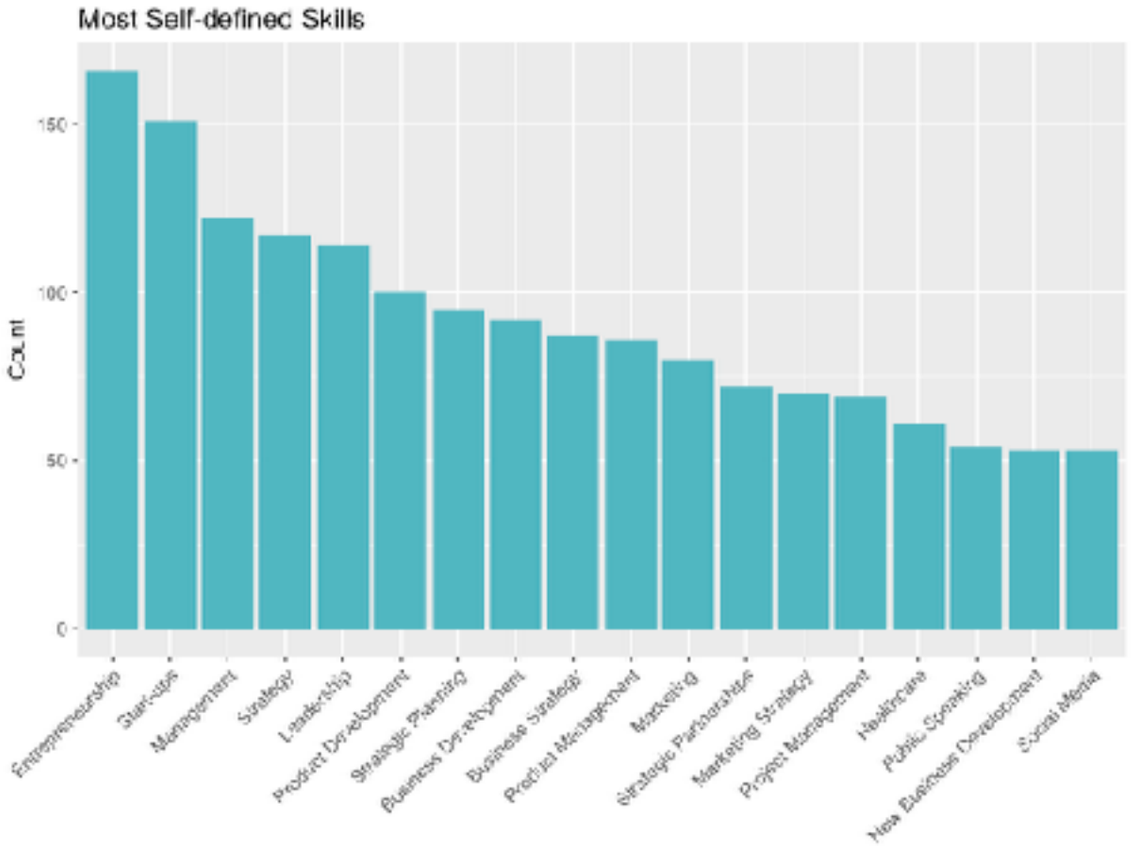


Likely to have former leading experience

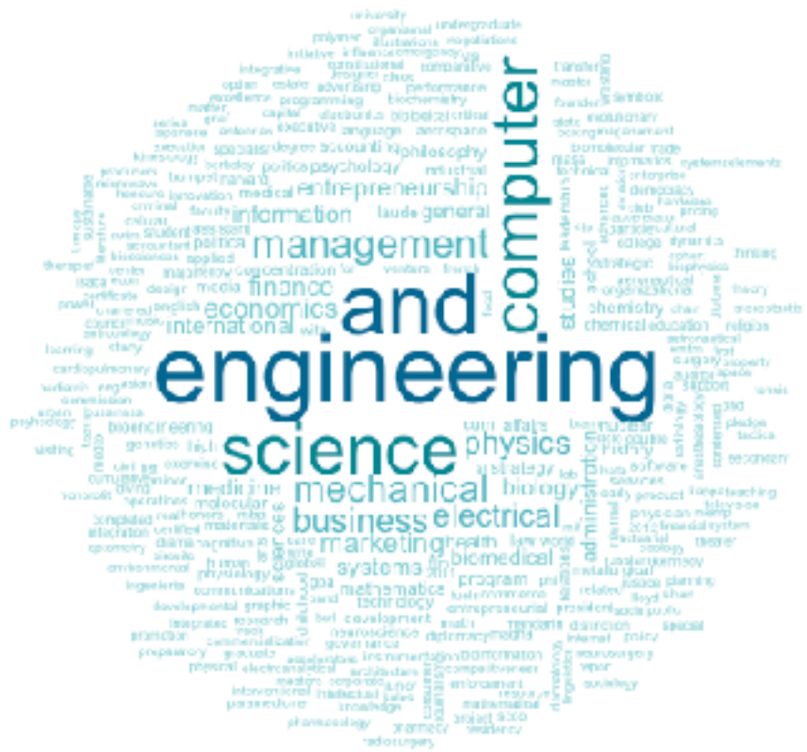


Q3: WHAT ARE THEY GOOD AT?

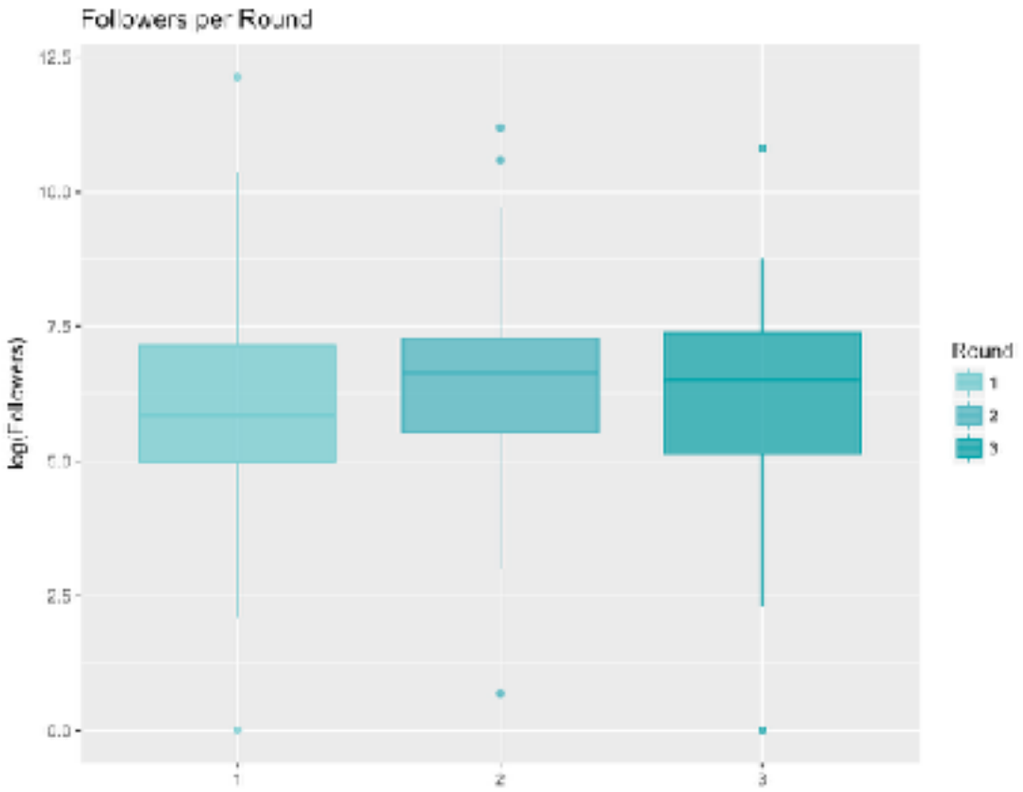
Management and leadership skills

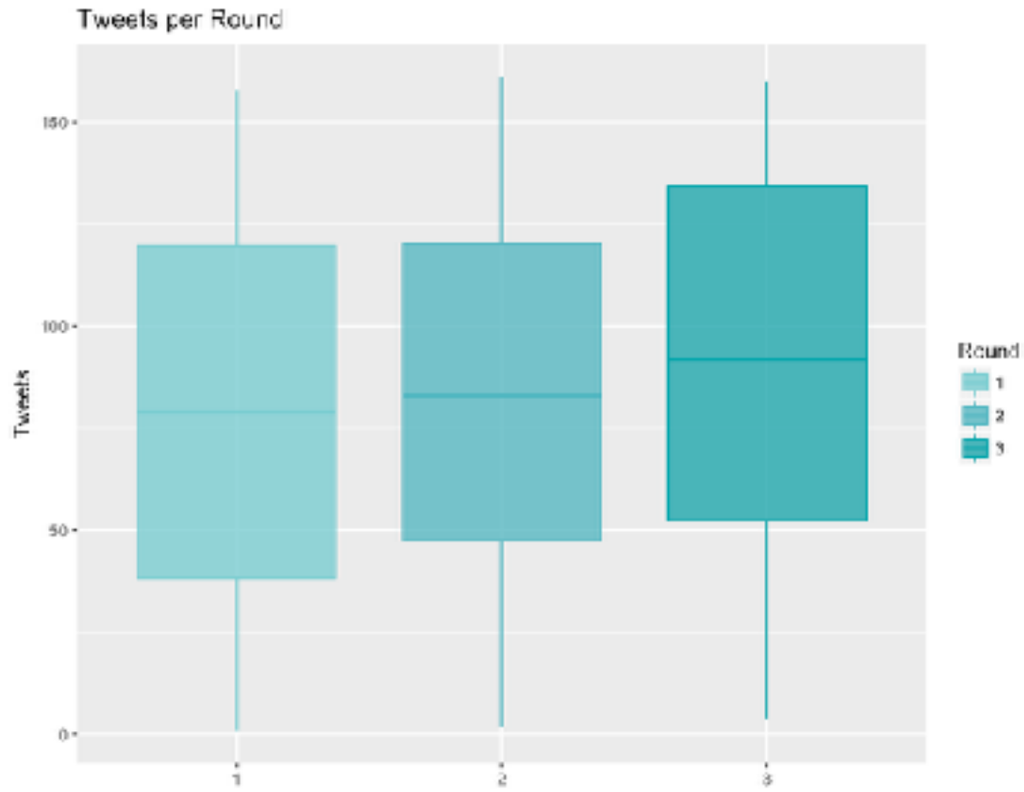


Science and engineering background



Q4: SOCIAL MEDIA VS FUNDING ROUNDS?





Q4: SOCIAL MEDIA VS FUNDING ROUNDS?