DSC324/424
Assignment #3 (Due Sunday, August 14, 2022 at midnight)

1) (20 points, for data projects) Choose a technique that we have covered so far in this course, and try applying that technique to your data. You may choose any of

a) Model building and Multiple Regression
b) PCA
c) CFA
d) CCA
e) CA (correspondence analysis)

If you are working as a group, each member of your group should try a different technique, or the same technique with different aspects of the data.

I performed the CCA technique on our group's data (Online New popularity data set). Our data is not really categorical data except for some variables on weekdays. Our dependent variable only had one variable so when creating the categorical variables seen in the code below we only had 1 CV. This limits our data and the information we are going to receive. For reference see Helio plots below the code.

```
library(yacca)


#Read in Data
```

```
setwd("C:/Users/doret/Documents/DSC 424/Project")


ONP = read.csv("OnlineNewsPopularity.csv", header = TRUE, sep = ",")
head(ONP)


#See the first six lines of the data
head(ONP)


names(ONP)


shares = ONP[, 61]
numbers = ONP[, 3:13]
data = ONP[, 14:19]
keyword = ONP[, 20:28]
selfRef = ONP[, 29:31]
day = ONP[, 32:39]
```

```
LDA = ONP[, 40:44]

global = ONP[, 45:50]

polarity = ONP[, 51:56]

tital = ONP[, 57:60]


#Numbers

# This gives us the cannonical correlates, but no significance tests

# c = cca(shares,numbers)

# summary(c)


#CV1

# helio.plot(c, cv=1, x.name="shares Values",

        y.name="numbers Values")


#Function Names

# ls(c)
```

```
# Perform a chi-square test on C

# c

# ls(c)

# c$chisq

# c$df

# summary(c)

# round(pchisq(c$chisq, c$df, lower.tail=F), 3)


#Data

# This gives us the cannonical correlates, but no significance tests

c2 = cca(shares,data)

summary(c2)


#CV1

helio.plot(c2, cv=1, x.name="shares Values",
```

```
        y.name="Data Values")


#Function Names
ls(c2)


# Perform a chi-square test on C2
c2
ls(c2)
c2$chisq
c2$df
summary(c2)
round(pchisq(c2$chisq, c2$df, lower.tail=F), 3)


#Keywords
# This gives us the cannonical correlates, but no significance tests
c3 = cca(shares,keyword)
```

```
summary(c3)


#CV1
helio.plot(c3, cv=1, x.name="shares Values",
          y.name="keyword Values")


#Function Names
ls(c3)


# Perform a chi-square test on C2
c3
ls(c3)
c3$chisq
c3$df
summary(c3)
round(pchisq(c3$chisq, c3$df, lower.tail=F), 3)
```

```
#selfRef
# This gives us the cannonical correlates, but no significance tests
c4 = cca(shares,selfRef)
summary(c4)

#CV1
helio.plot(c4, cv=1, x.name="shares Values",
        y.name="selfRef Values")

#Function Names
ls(c4)

# Perform a chi-square test on C2
c4
ls(c4)
```

```
c4$chisq

c4$df

summary(c4)

round(pchisq(c4$chisq, c4$df, lower.tail=F), 3)


# #day

# # This gives us the cannonical correlates, but no significance tests

# c5 = cca(shares,day)

# summary(c5)

#

# #CV1

# helio.plot(c5, cv=1, x.name="shares Values",

#            y.name="day Values")

#

# #Function Names

# ls(c5)
```

```
#

# # Perform a chi-square test on C2

# c5

# ls(c5)

# c5$chisq

# c5$df

# summary(c5)

# round(pchisq(c5$chisq, c5$df, lower.tail=F), 3)


#LDA

# This gives us the cannonical correlates, but no significance tests

c6 = cca(shares,LDA)

summary(c6)


#CV1

helio.plot(c6, cv=1, x.name="shares Values",
```

```
        y.name="LDA Values")


#Function Names
ls(c6)


# Perform a chi-square test on C2
c6
ls(c6)
c6$chisq
c6$df
summary(c6)
round(pchisq(c6$chisq, c6$df, lower.tail=F), 3)


#global
# This gives us the cannonical correlates, but no significance tests
c7 = cca(shares,global)
```

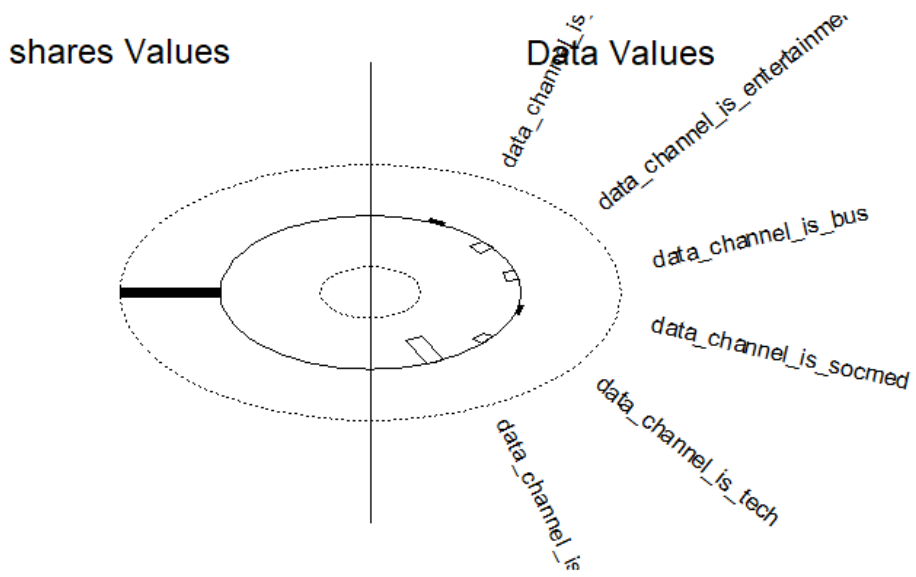
```
summary(c7)


#CV1
helio.plot(c7, cv=1, x.name="shares Values",
          y.name="global Values")


#Function Names
ls(c7)


# Perform a chi-square test on C2
c7
ls(c7)
c7$chisq
c7$df
summary(c7)
round(pchisq(c7$chisq, c7$df, lower.tail=F), 3)
```

```
#polarity
# This gives us the cannonical correlates, but no significance tests
c8 = cca(shares,polarity)
summary(c8)

#CV1
helio.plot(c8, cv=1, x.name="shares Values",
        y.name="polarity Values")

#Function Names
ls(c8)

# Perform a chi-square test on C2
c8
ls(c8)
```

```
c8$chisq

c8$df

summary(c8)

round(pchisq(c8$chisq, c8$df, lower.tail=F), 3)


#tital

# This gives us the cannonical correlates, but no significance tests

c9 = cca(shares, tital)

summary(c9)


#CV1

helio.plot(c9, cv=1, x.name="shares Values",

        y.name="tital Values")


#Function Names

ls(c9)
```

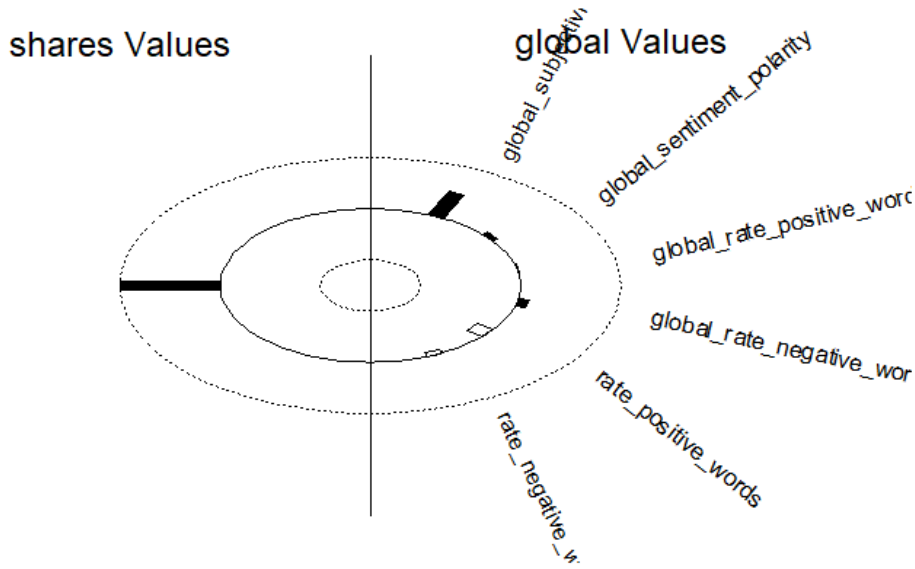
```
# Perform a chi-square test on C2

c9

ls(c9)

c9$chisq

c9$df

summary(c9)

round(pchisq(c9$chisq, c9$df, lower.tail=F), 3)
```

Helio Plot



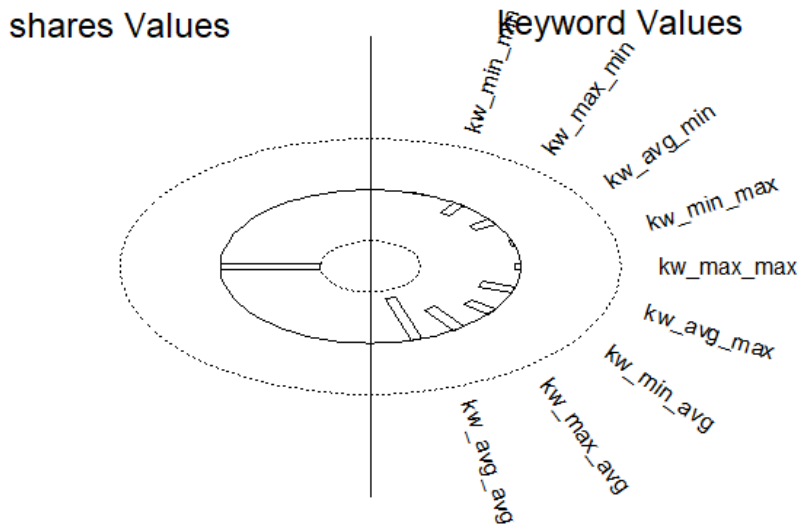
Canonical Variate1

Helio Plot



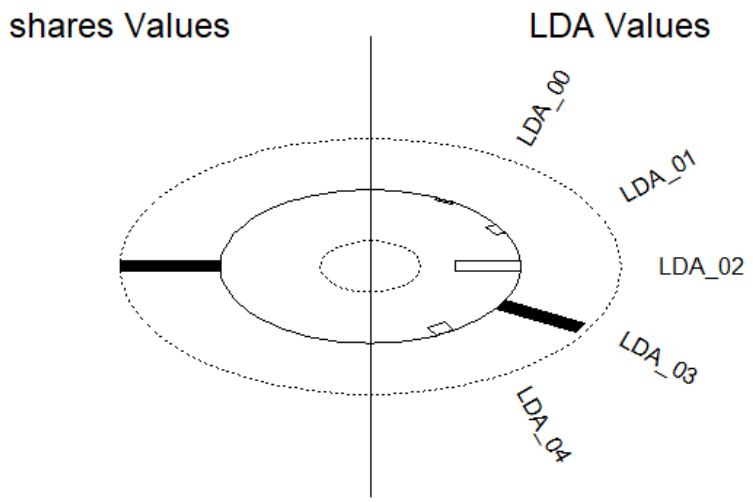
Canonical Variate1

Helio Plot



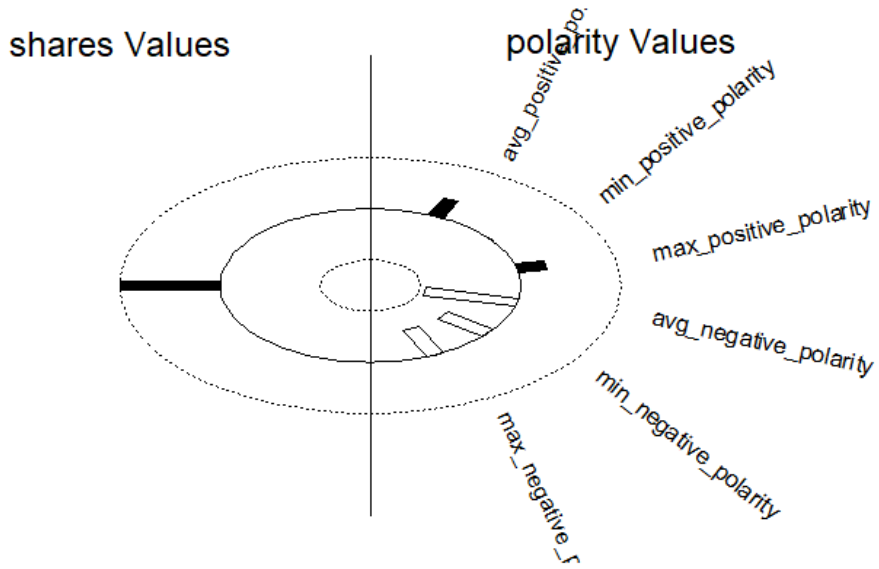
Canonical Variate1

Helio Plot

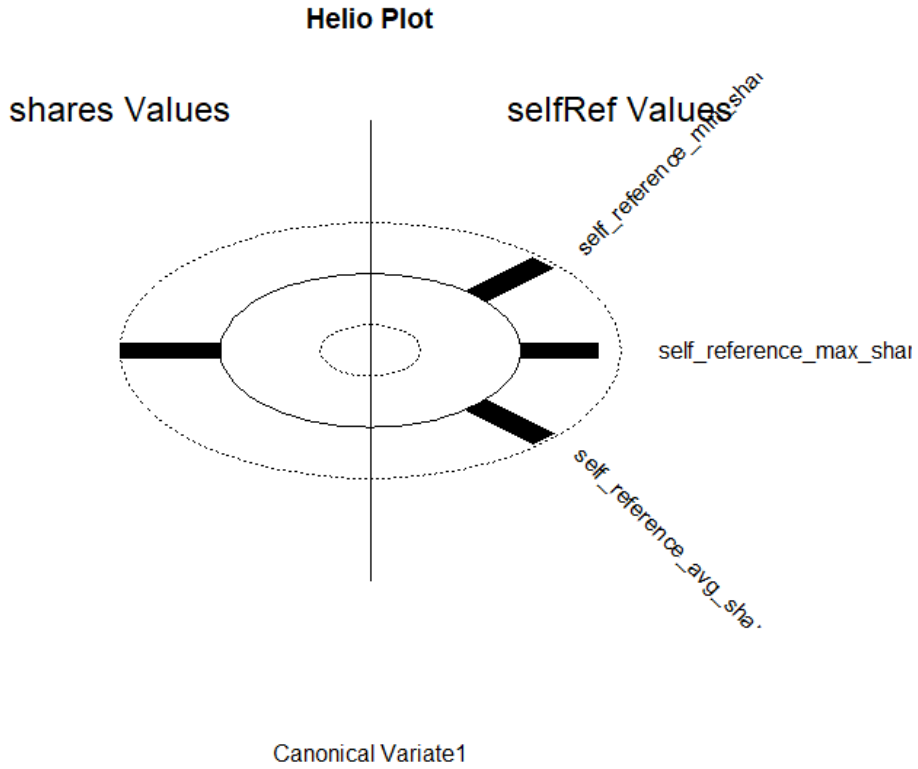


Canonical Variate1

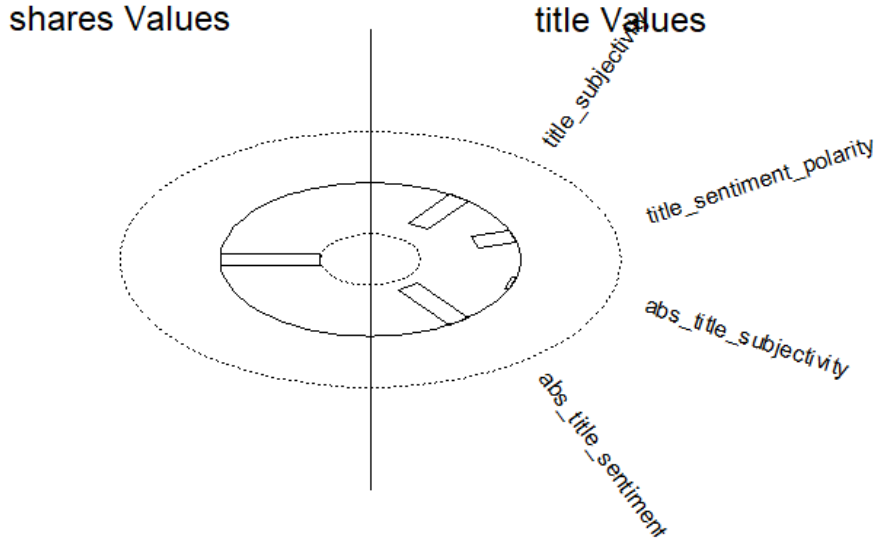
Helio Plot



Canonical Variate1


Helio Plot
shares Values
selfRef Values
self_reference_max_shar
Canonical Variate1

Helio Plot



Canonical Variate1