Analysis Component Group E

E. Hauge, G. Henry, L. Kaminka

2024-03-28

```
NewYork_SanFrancisco_data <- read_csv("data/NewYork_SanFrancisco_data.csv", show_col_types = FALSE)
```

```
## New names:
## * `` -> `...1`
```

```
names(NewYork_SanFrancisco_data)
```

```
##   [1] "...1"          "OBJECTID"      "GEOID10"       "GEOID20"       "STATEFP"
##   [6] "COUNTYFP"      "TRACTCE"       "BLKGRPCE"      "CSA"           "CSA_Name"
##  [11] "CBSA"          "CBSA_Name"     "CBSA_POP"      "CBSA_EMP"      "CBSA_WRK"
##  [16] "Ac_Total"      "Ac_Water"      "Ac_Land"       "Ac_Unpr"       "TotPop"
##  [21] "CountHU"       "HH"            "P_WrkAge"      "AutoOwn0"      "Pct_AO0"
##  [26] "AutoOwn1"      "Pct_AO1"       "AutoOwn2p"     "Pct_AO2p"      "Workers"
##  [31] "R_LowWageWk"   "R_MedWageWk"   "R_HiWageWk"    "R_PCTLOWWAGE"  "TotEmp"
##  [36] "E5_Ret"        "E5_Off"        "E5_Ind"        "E5_Svc"        "E5_Ent"
##  [41] "E8_Ret"        "E8_off"        "E8_Ind"        "E8_Svc"        "E8_Ent"
##  [46] "E8_Ed"         "E8_Hlth"       "E8_Pub"        "E_LowWageWk"   "E_MedWageWk"
##  [51] "E_HiWageWk"    "E_PctLowWage"  "D1A"           "D1B"           "D1C"
##  [56] "D1C5_RET"      "D1C5_OFF"      "D1C5_IND"      "D1C5_SVC"      "D1C5_ENT"
##  [61] "D1C8_RET"      "D1C8_OFF"      "D1C8_IND"      "D1C8_SVC"      "D1C8_ENT"
##  [66] "D1C8_ED"       "D1C8_HLTH"     "D1C8_PUB"      "D1D"           "D1_FLAG"
##  [71] "D2A_JPHH"      "D2B_E5MIX"     "D2B_E5MIXA"    "D2B_E8MIX"     "D2B_E8MIXA"
##  [76] "D2A_EPHHM"     "D2C_TRPMX1"    "D2C_TRPMX2"    "D2C_TRIPEQ"    "D2R_JOBPOP"
##  [81] "D2R_WRKEMP"    "D2A_WRKEMP"    "D2C_WREMLX"    "D3A"           "D3AAO"
##  [86] "D3AMM"         "D3APO"         "D3B"           "D3BAO"         "D3BMM3"
##  [91] "D3BMM4"        "D3BPO3"        "D3BPO4"        "D4A"           "D4B025"
##  [96] "D4B050"        "D4C"           "D4D"           "D4E"           "D5AR"
## [101] "D5AE"          "D5BR"          "D5BE"          "D5CR"          "D5CRI"
## [106] "D5CE"          "D5CEI"         "D5DR"          "D5DRI"         "D5DE"
## [111] "D5DEI"         "D2A_Ranked"    "D2B_Ranked"    "D3B_Ranked"    "D4A_Ranked"
## [116] "NatWalkInd"    "Shape_Length"  "Shape_Area"
```

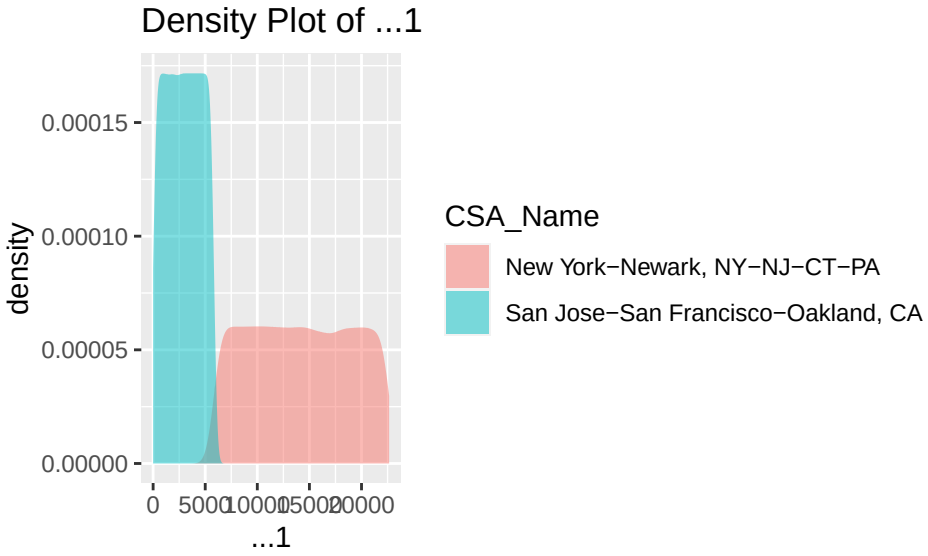
```
#removing missing values.
NewYork_SanFrancisco_data <- na.omit(NewYork_SanFrancisco_data)
```

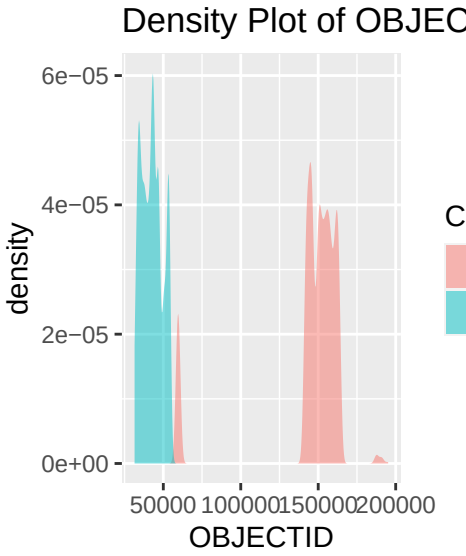
#Univariate Analysis:

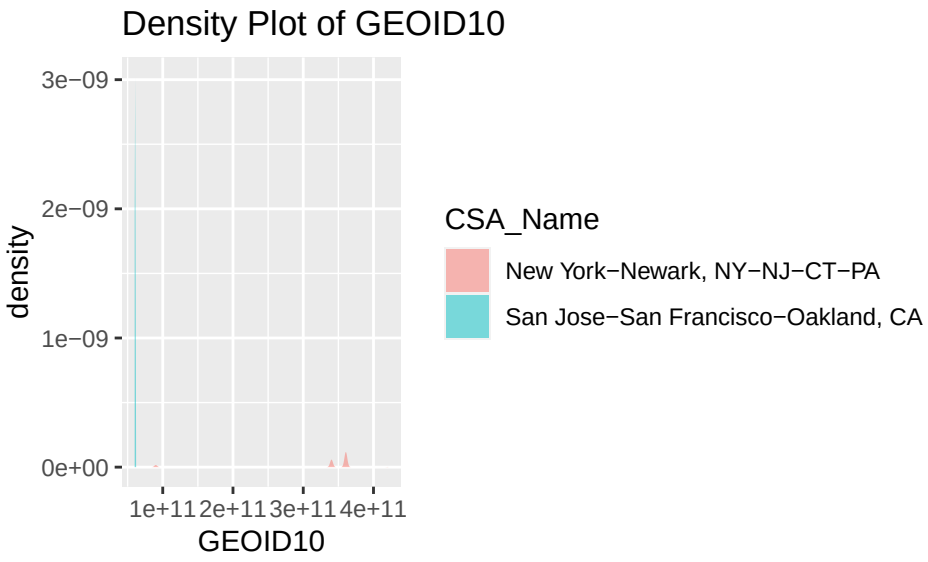
```
create_density_plots_subset <- function(data, start_col, end_col) {
  numeric_vars <- names(data)[sapply(data, is.numeric)]
  selected_vars <- numeric_vars[start_col:end_col]
```

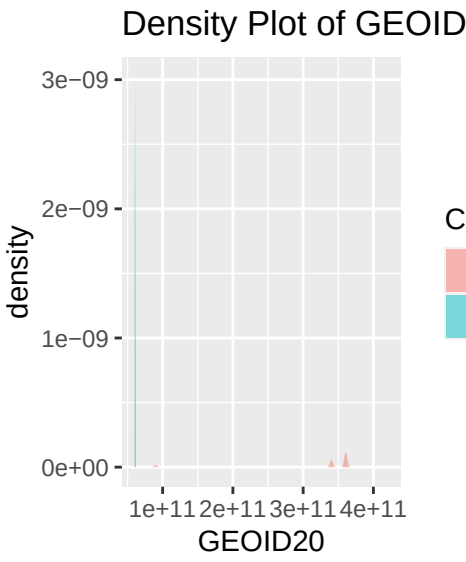
```
  for (var in selected_vars) {
    # Create density plot for each variable
    plot <- gf_density(as.formula(paste("~", var)), data = data, fill = ~CSA_Name,
                       title = paste("Density Plot of", var))
    print(plot)
  }
}

# Function to create density plots for all numeric variables


# Example usage:
# Assuming 'data' is your dataset and you want to create density plots for columns 1 to 10:
```

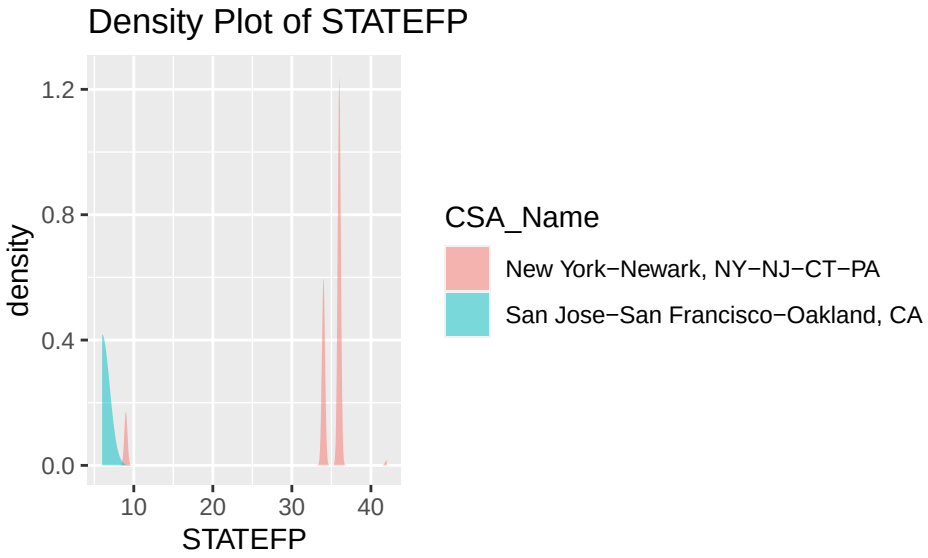
```
create_density_plots_subset(NewYork_SanFrancisco_data, 1, 25)
```

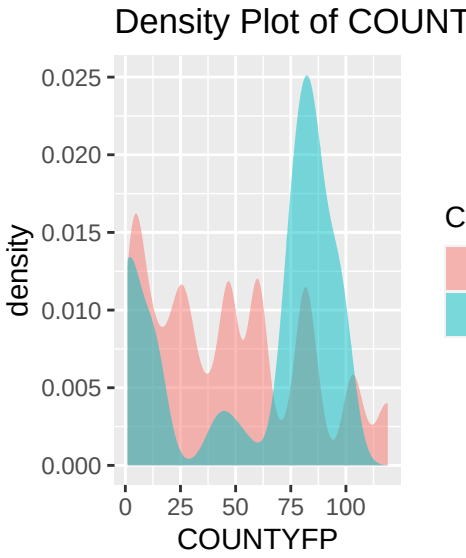


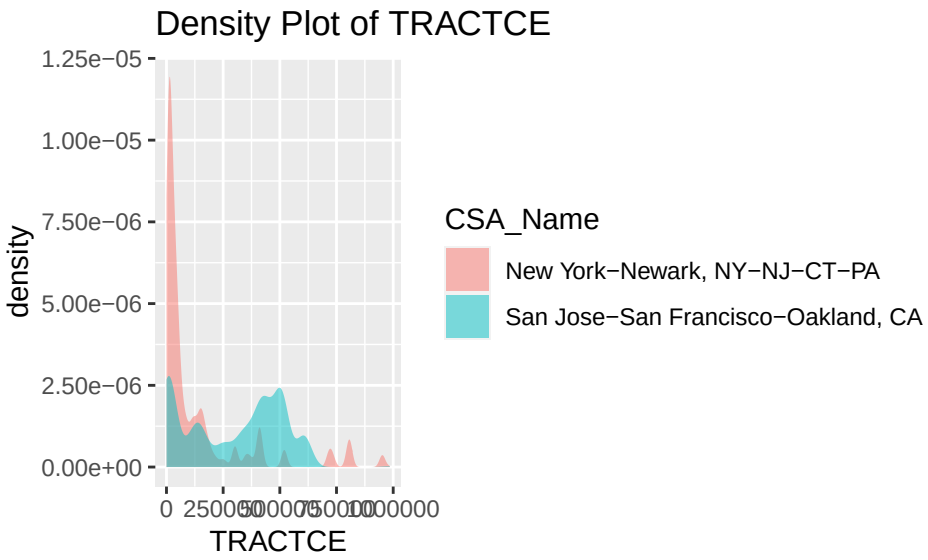


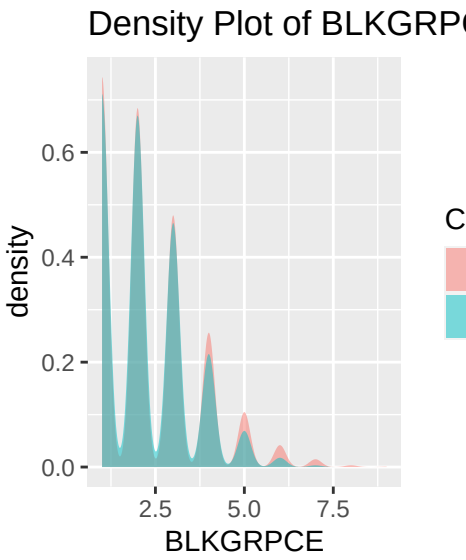


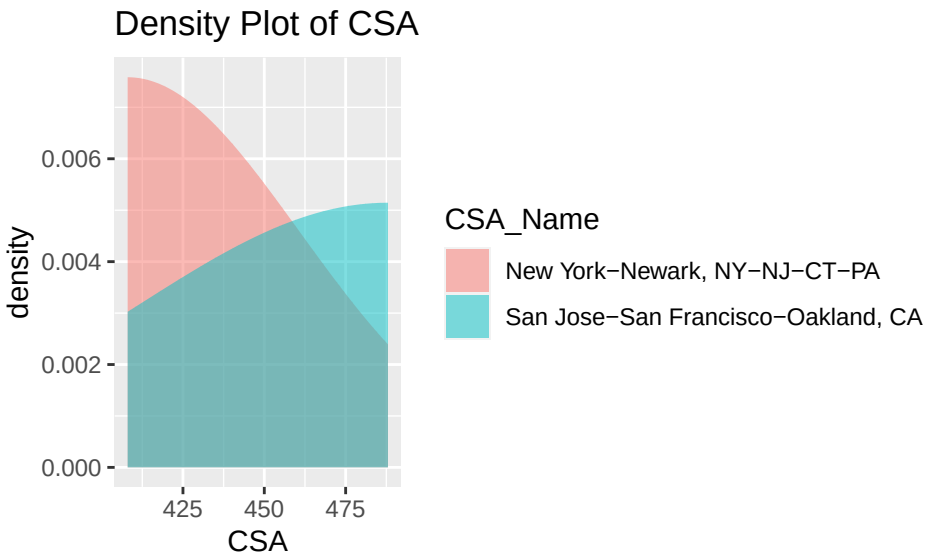


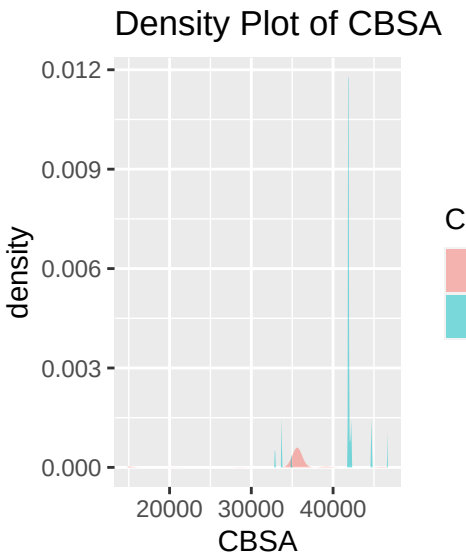

Density Plot of STATEFP
density
1.2
0.8
0.4
0.0
10
20
30
40
STATEFP
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

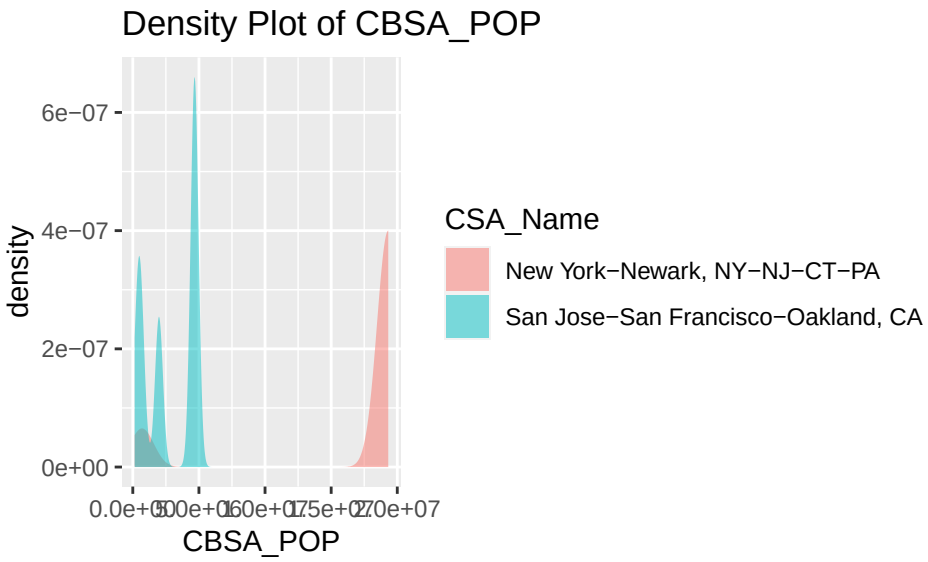

Density Plot of COUNT
density
0.025
0.020
0.015
0.010
0.005
0.000
0
25
50
75
100
COUNTYFP

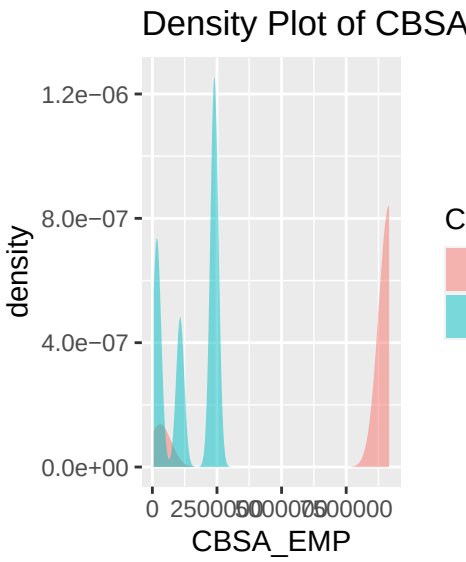

Density Plot of TRACTCE
density
1.25e−05
1.00e−05
7.50e−06
5.00e−06
2.50e−06
0.00e+00
0
TRACTCE
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

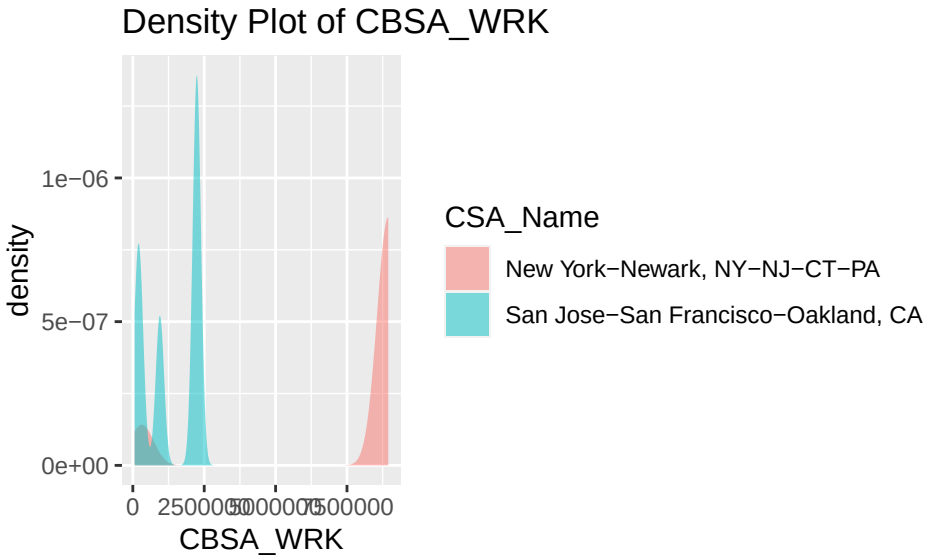

Density Plot of BLKGRP
density
0.6
0.4
0.2
0.0
2.5
5.0
7.5
BLKGRPCE

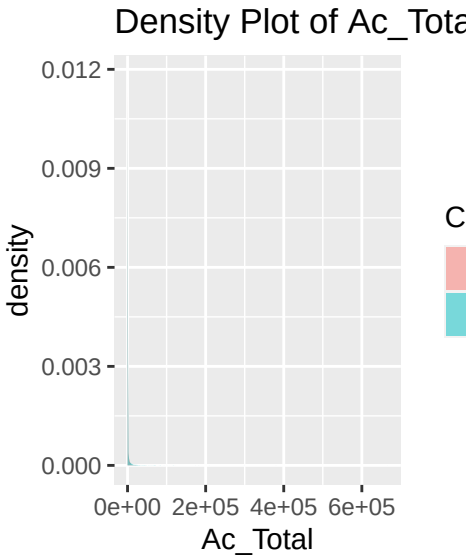

Density Plot of CSA
density
0.006
0.004
0.002
0.000
425
450
475
CSA
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

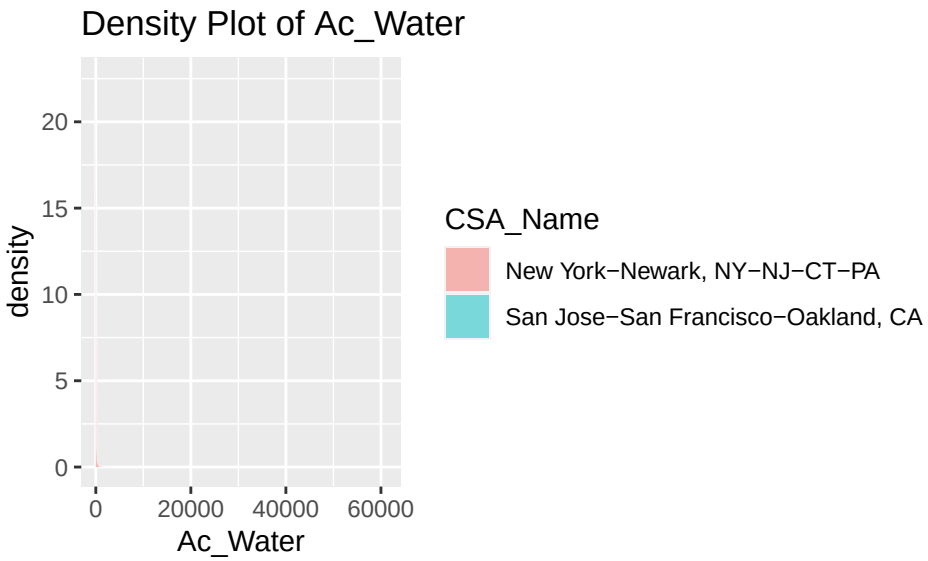

Density Plot of CBSA
density
0.012
0.009
0.006
0.003
0.000
20000
30000
40000
CBSA

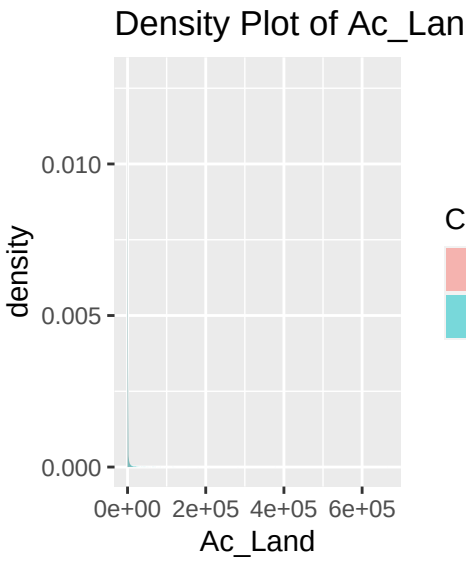

Density Plot of CBSA_POP
density
6e−07
4e−07
2e−07
0e+00
CBSA_POP
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

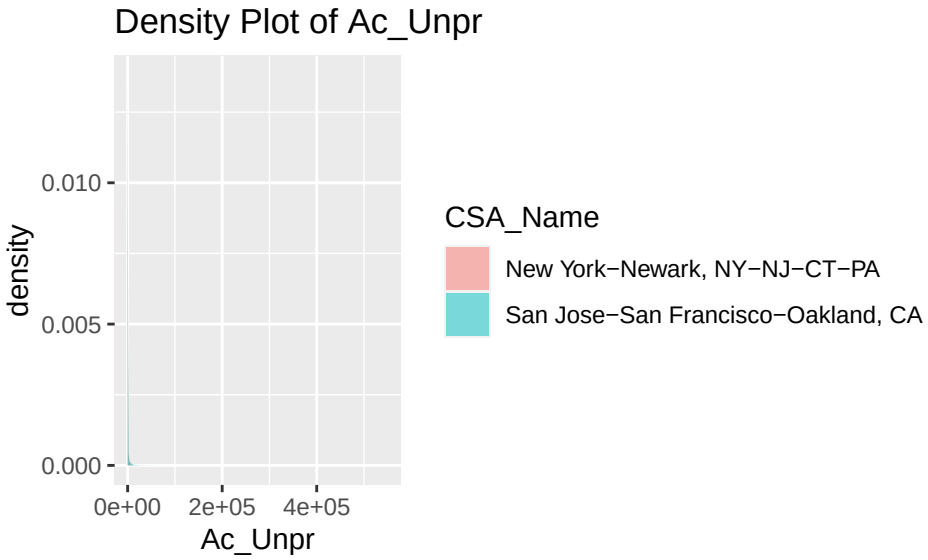

Density Plot of CBSA
density
1.2e−06
8.0e−07
4.0e−07
0.0e+00
CBSA_EMP

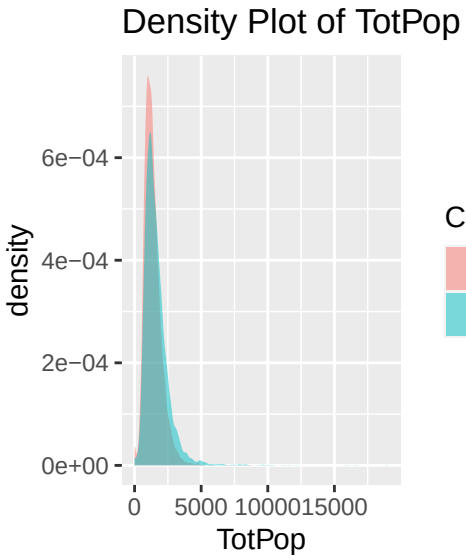

Density Plot of CBSA_WRK
density
1e−06
5e−07
0e+00
0
CBSA_WRK
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

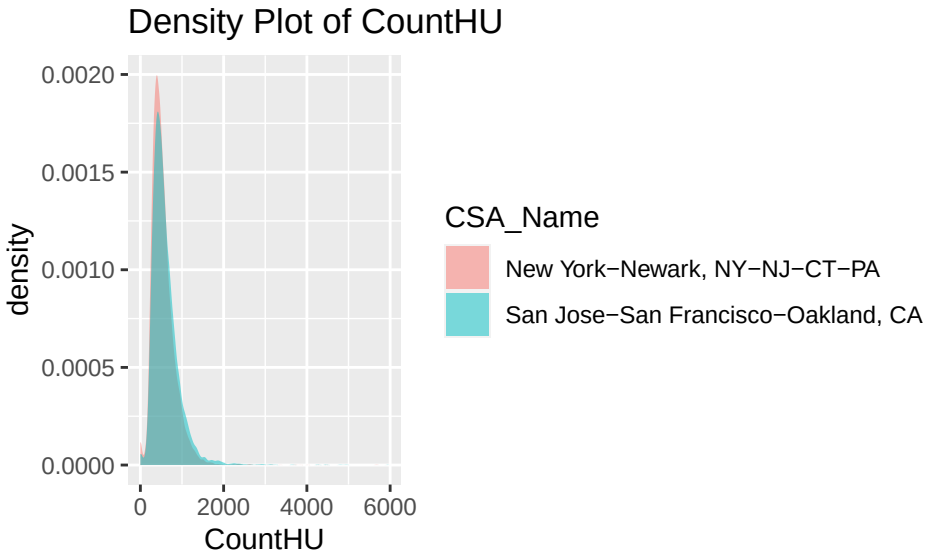

Density Plot of Ac_Tota
density
0.012
0.009
0.006
0.003
0.000
0e+00
2e+05
4e+05
6e+05
Ac_Total
C

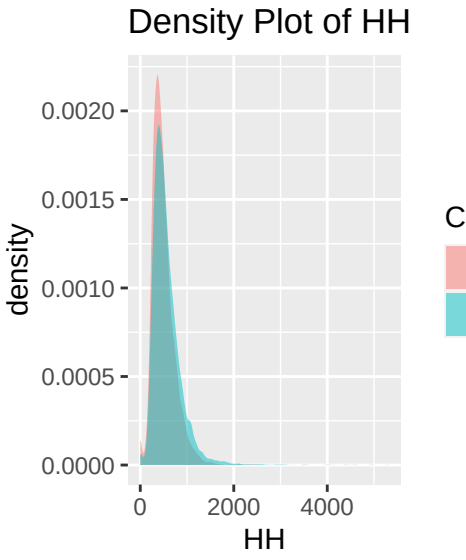

Density Plot of Ac_Water
density
20
15
10
5
0
0
20000
40000
60000
Ac_Water
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

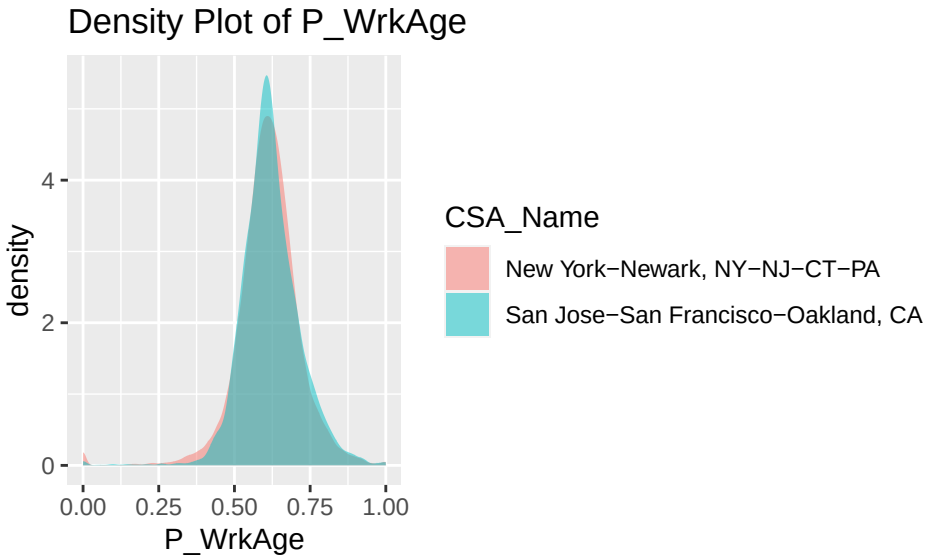

Density Plot of Ac_Lan
density
0.010
0.005
0.000
0e+00
2e+05
4e+05
6e+05
Ac_Land
C

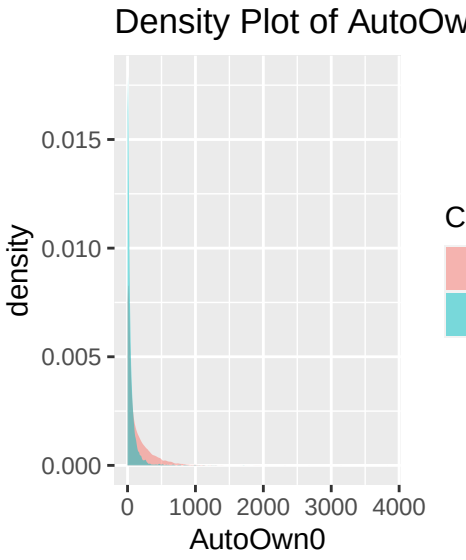

Density Plot of Ac_Unpr
density
0.010
0.005
0.000
0e+00
2e+05
4e+05
Ac_Unpr
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

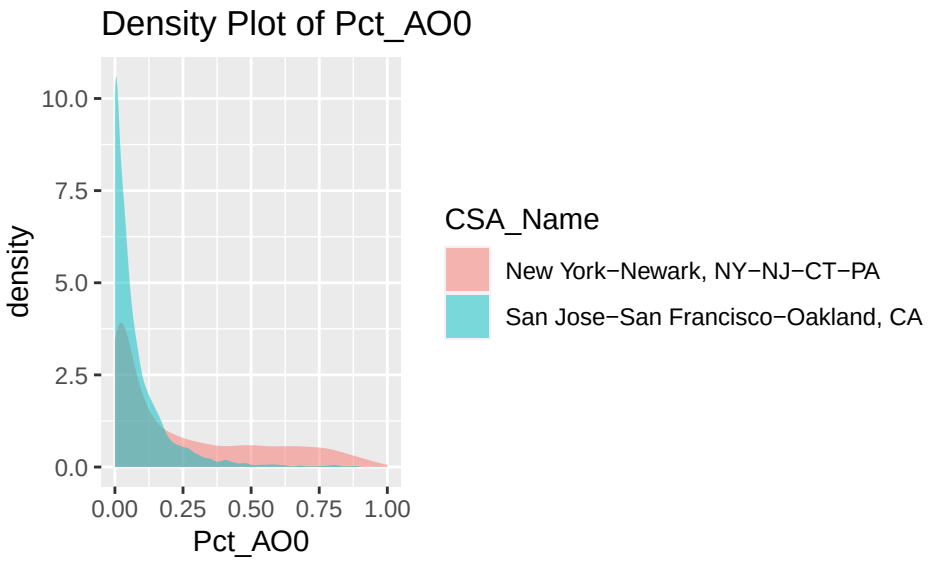

Density Plot of TotPop
density
6e−04
4e−04
2e−04
0e+00
0
5000
10000
15000
TotPop

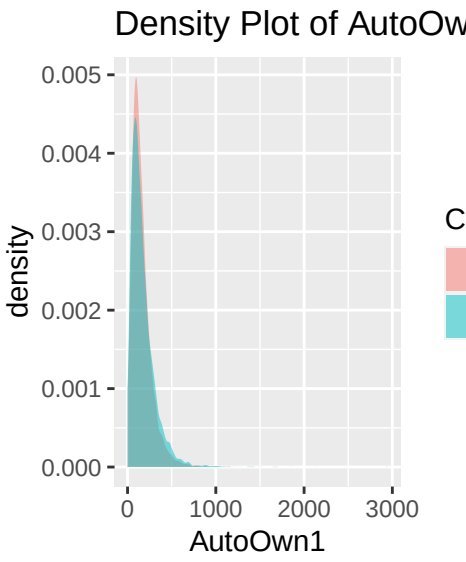

Density Plot of CountHU
density
0.0020
0.0015
0.0010
0.0005
0.0000
0
2000
4000
6000
CountHU
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

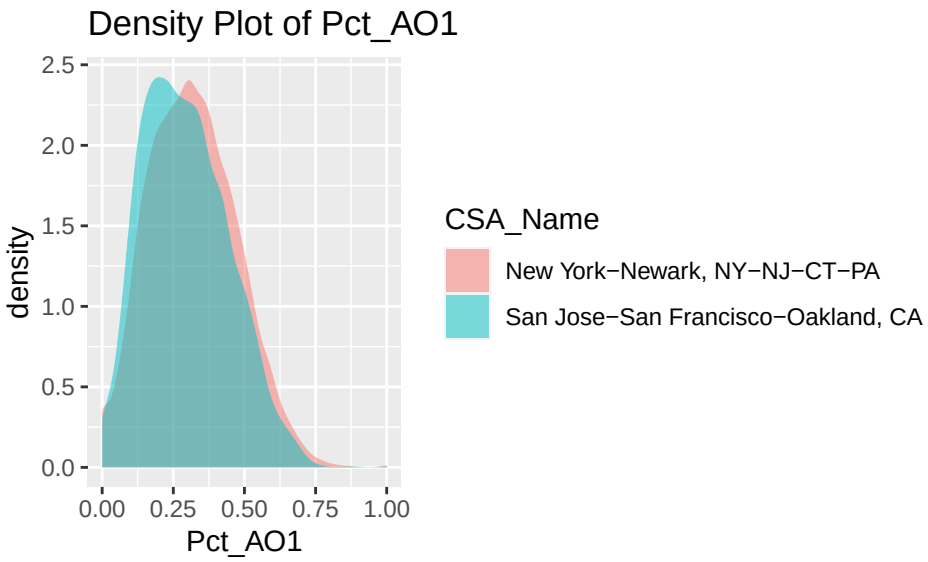

Density Plot of HH
density
0.0020
0.0015
0.0010
0.0005
0.0000
0
2000
4000
HH


Density Plot of P_WrkAge
density
P_WrkAge
0.00
0.25
0.50
0.75
1.00
0
2
4
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

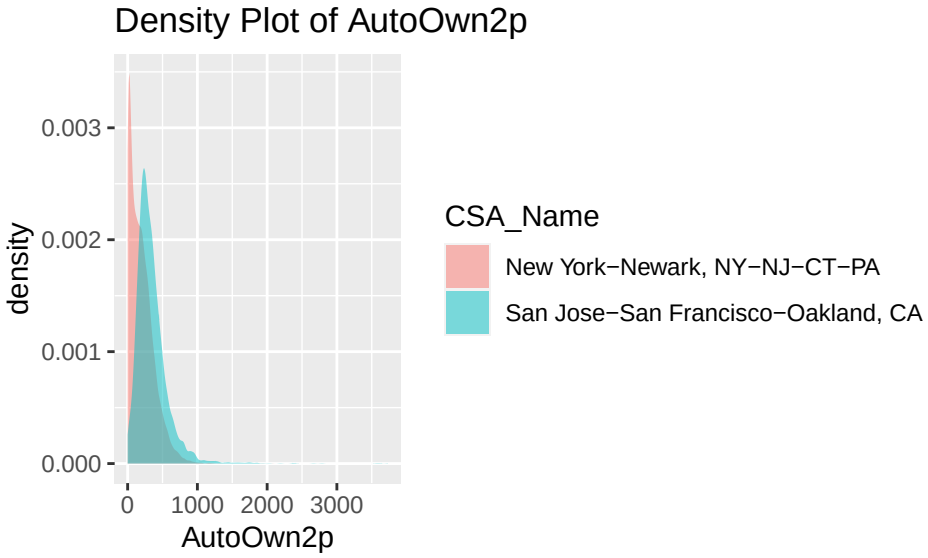

Density Plot of AutoOw
density
AutoOwn0
0
1000
2000
3000
4000
0.000
0.005
0.010
0.015

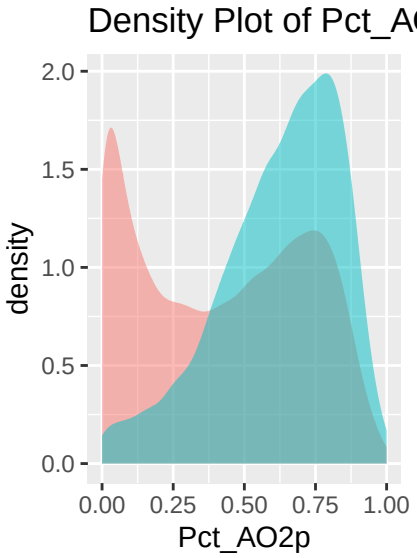

Density Plot of Pct_AO0
density
Pct_AO0
0.00
0.25
0.50
0.75
1.00
0.0
2.5
5.0
7.5
10.0
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

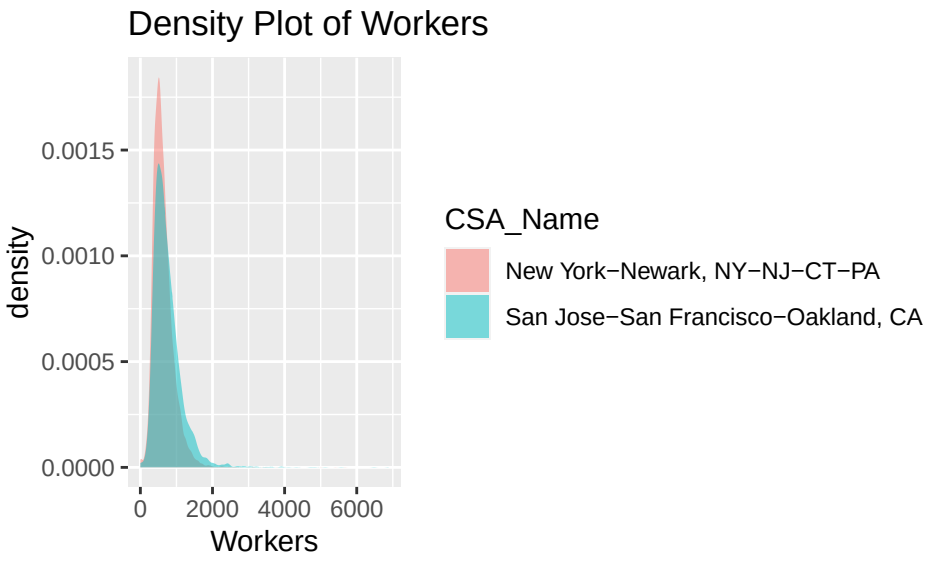

Density Plot of AutoOw
density
AutoOwn1
0
1000
2000
3000
0.000
0.001
0.002
0.003
0.004
0.005

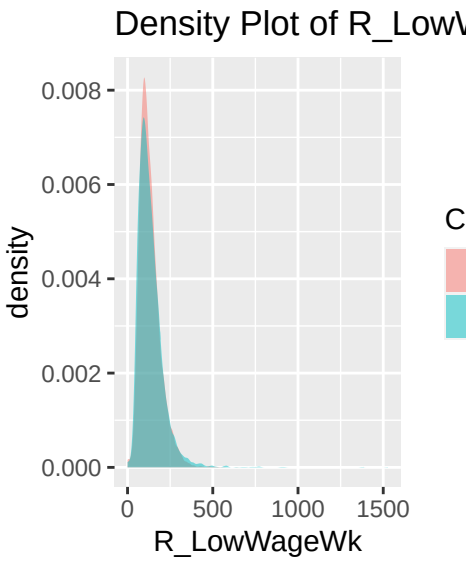


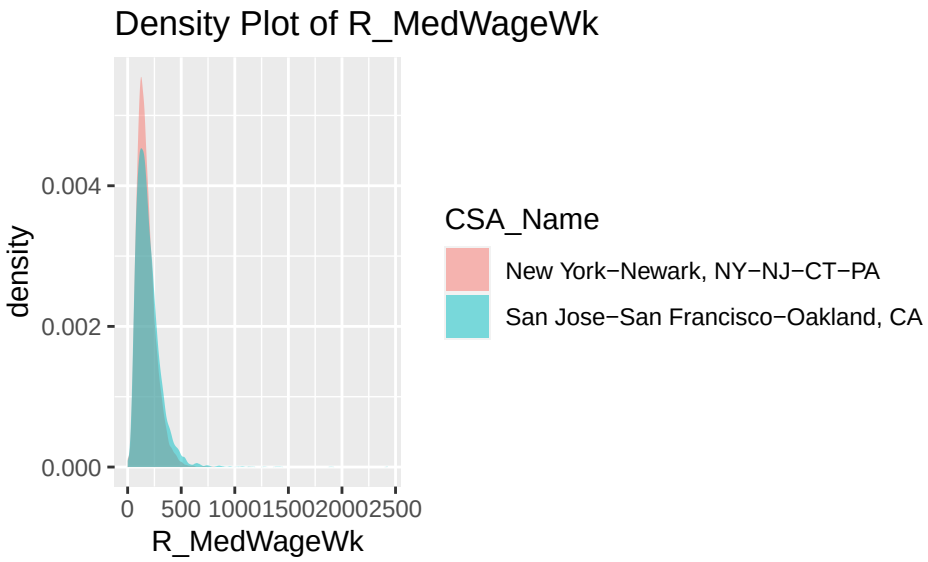

```
create_density_plots_subset(NewYork_SanFrancisco_data, 26, 51)
```

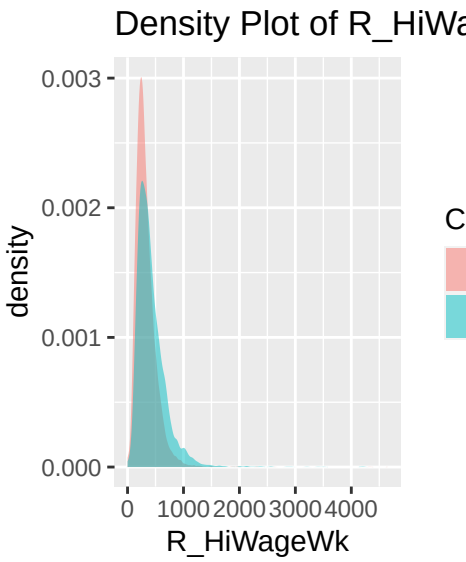


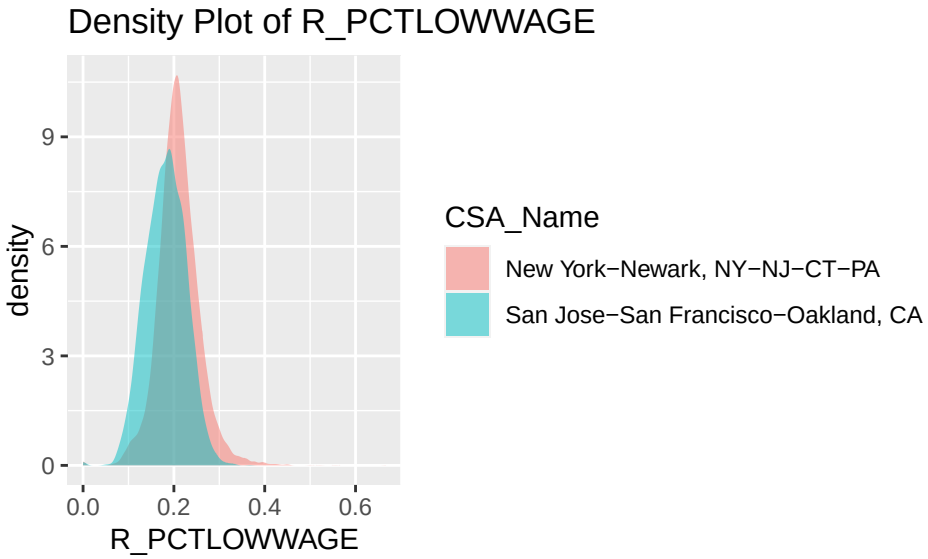


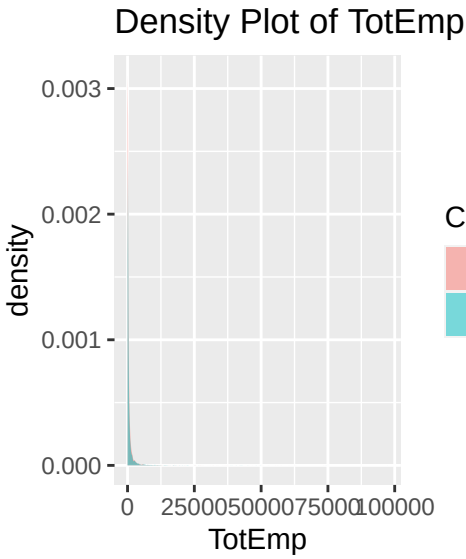

Density Plot of Workers
density
0.0015
0.0010
0.0005
0.0000
0
2000
4000
6000
Workers
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

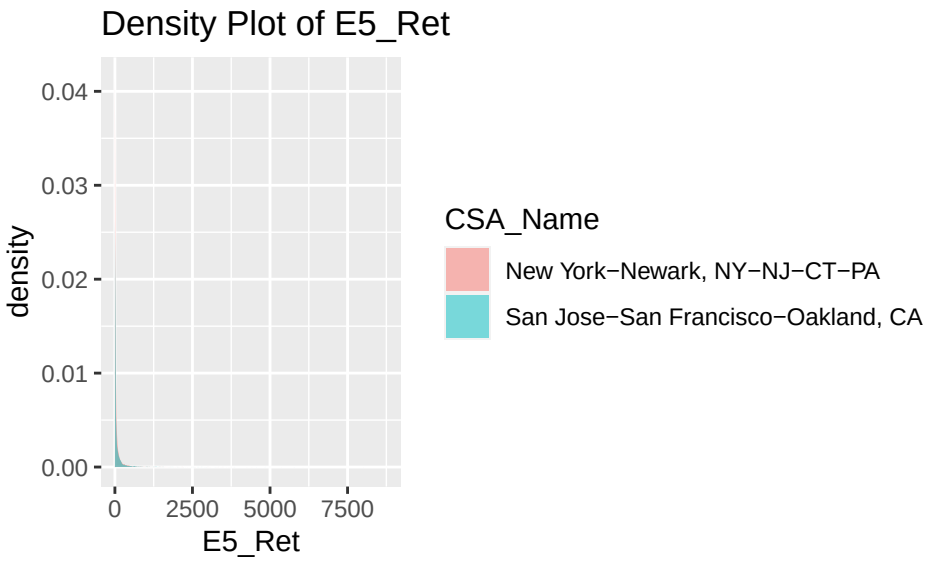

Density Plot of R_LowW
density
0.008
0.006
0.004
0.002
0.000
0
500
1000
1500
R_LowWageWk

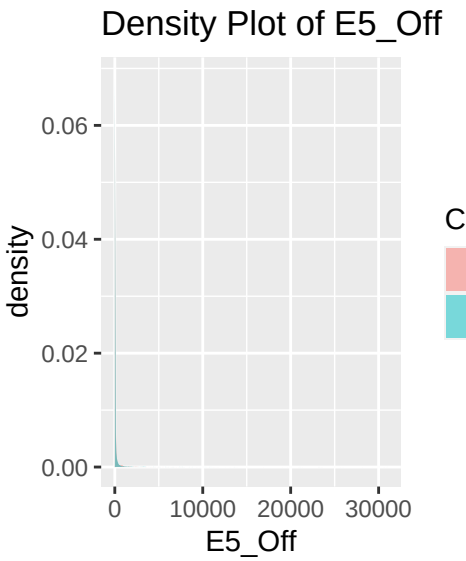

Density Plot of R_MedWageWk
density
0.004
0.002
0.000
0
500
1000
1500
2000
2500
R_MedWageWk
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

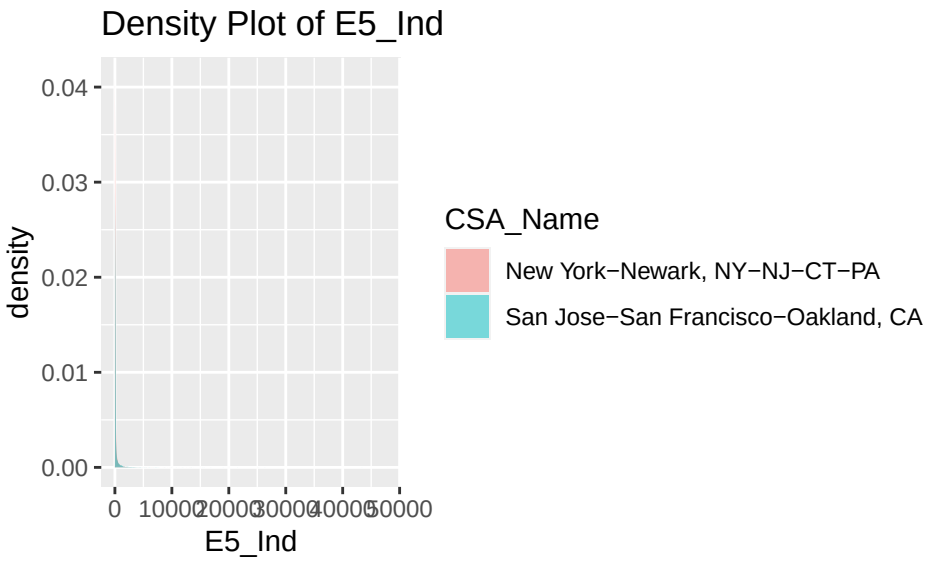

Density Plot of R_HiWa
density
0.003
0.002
0.001
0.000
0
1000
2000
3000
4000
R_HiWageWk

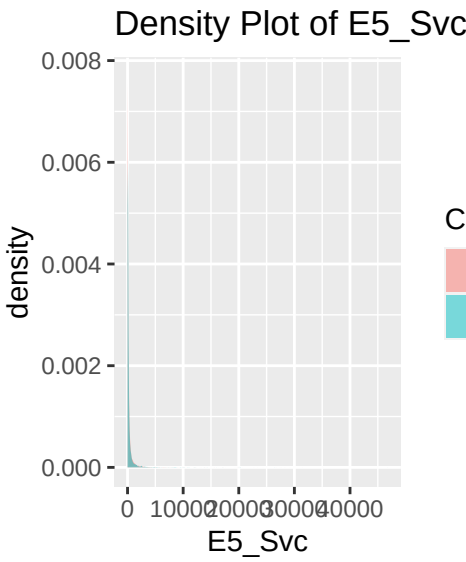

Density Plot of R_PCTLOWWAGE
density
R_PCTLOWWAGE
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

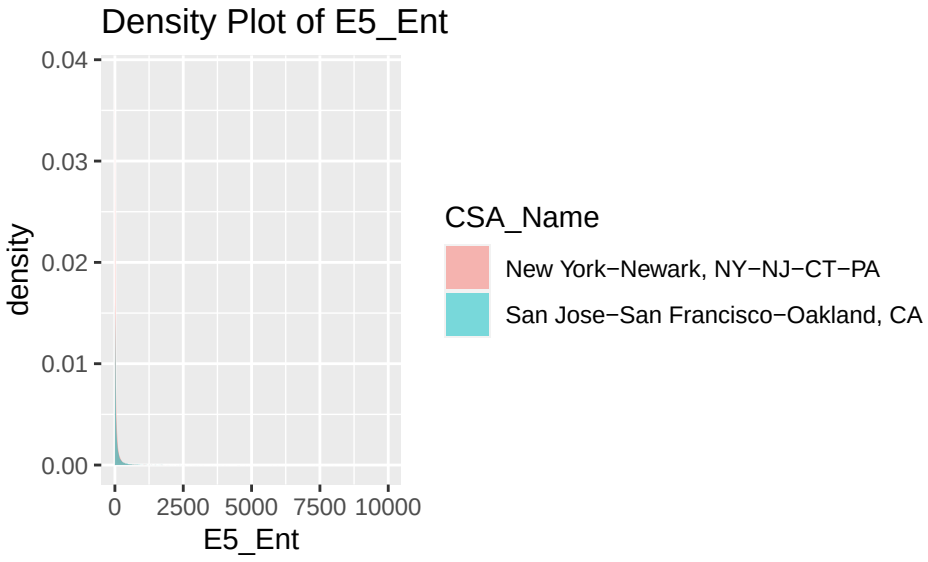

Density Plot of TotEmp
density
TotEmp

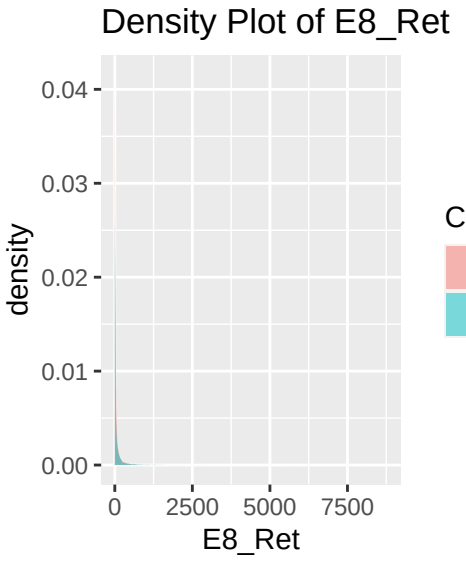

Density Plot of E5_Ret
density
E5_Ret
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

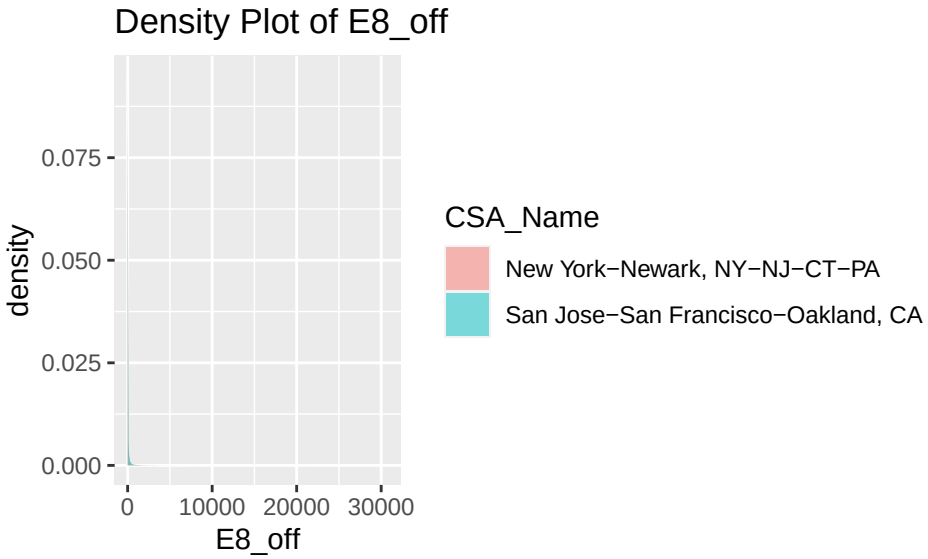

Density Plot of E5_Off
density
E5_Off

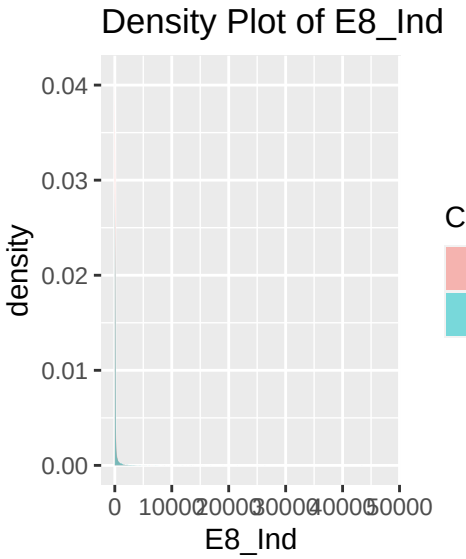

Density Plot of E5_Ind
density
0.04
0.03
0.02
0.01
0.00
0
10000
20000
30000
40000
50000
E5_Ind
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

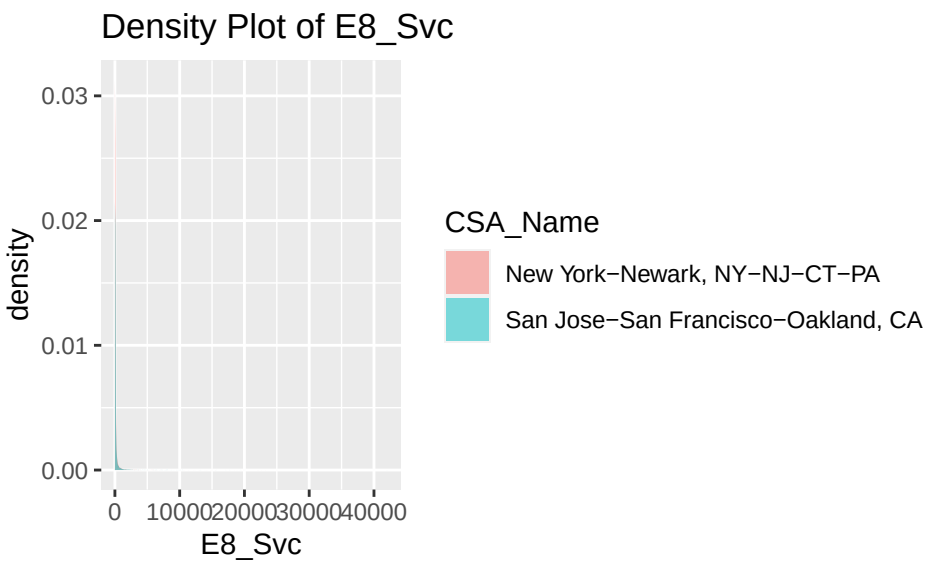

Density Plot of E5_Svc
density
0.008
0.006
0.004
0.002
0.000
0
10000
20000
30000
40000
E5_Svc
C

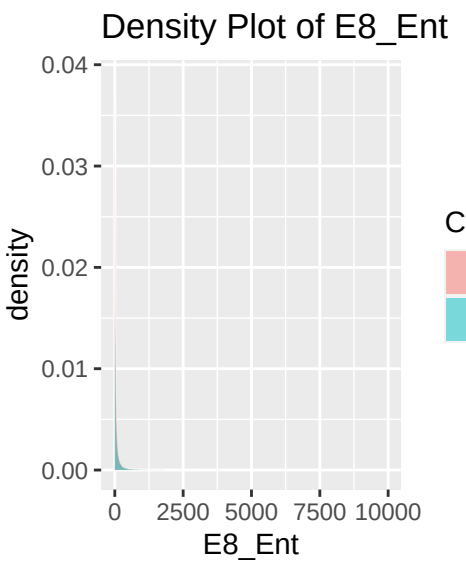

Density Plot of E5_Ent
density
0.04
0.03
0.02
0.01
0.00
0
2500
5000
7500
10000
E5_Ent
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

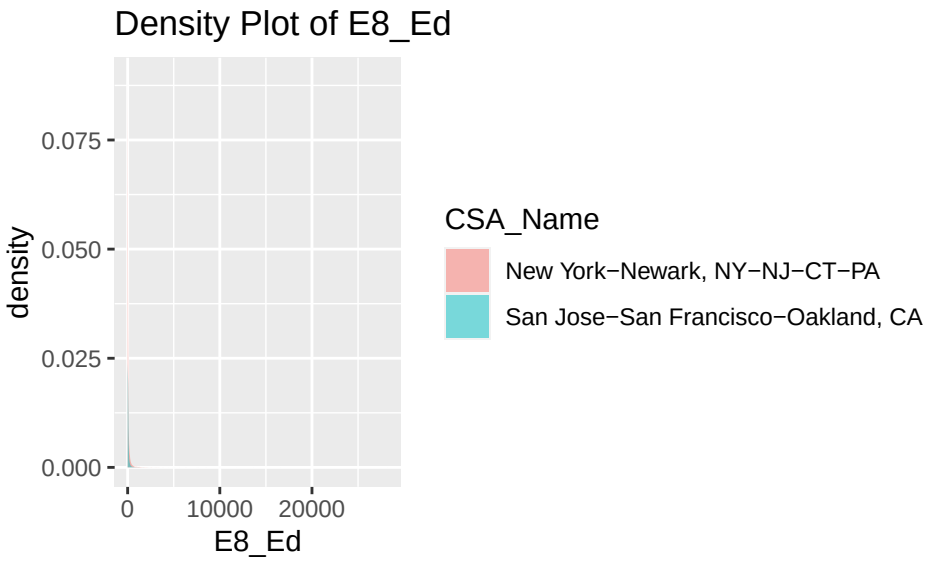

Density Plot of E8_Ret
density
0.04
0.03
0.02
0.01
0.00
0
2500
5000
7500
E8_Ret
C

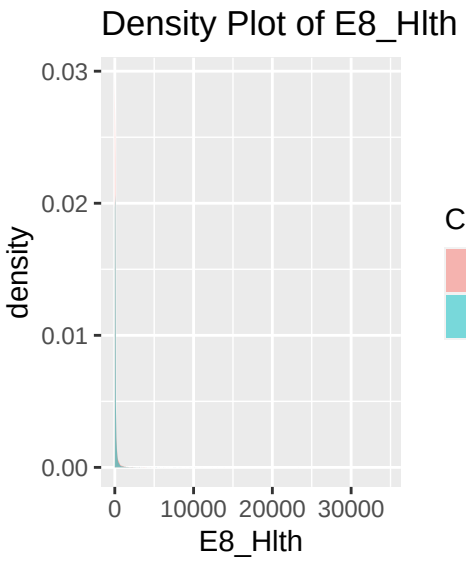

Density Plot of E8_off
density
0.075
0.050
0.025
0.000
0
10000
20000
30000
E8_off
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

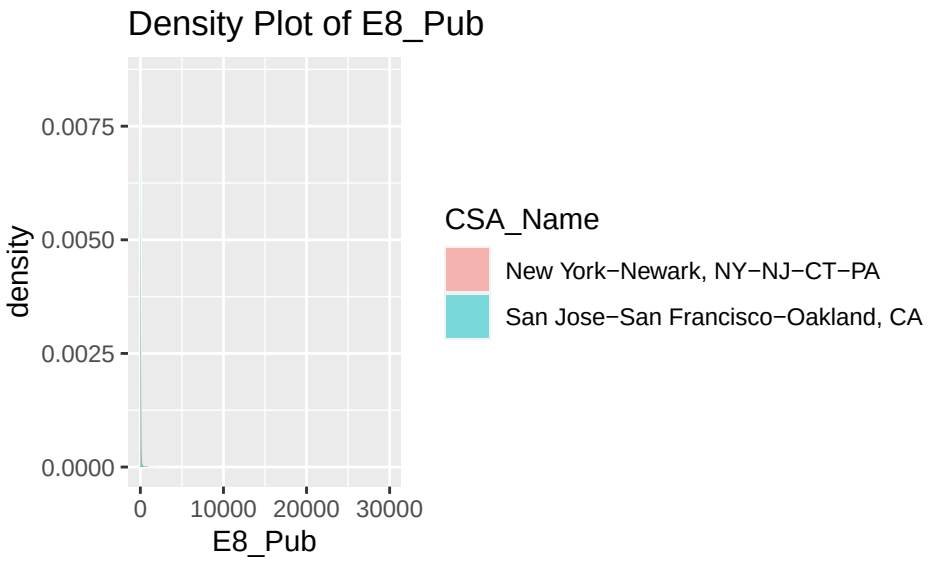

Density Plot of E8_Ind
density
0.04
0.03
0.02
0.01
0.00
0
E8_Ind

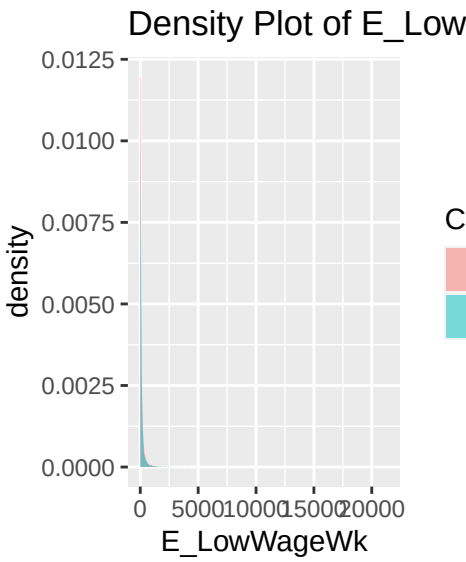

Density Plot of E8_Svc
density
0.03
0.02
0.01
0.00
0
10000
20000
30000
40000
E8_Svc
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

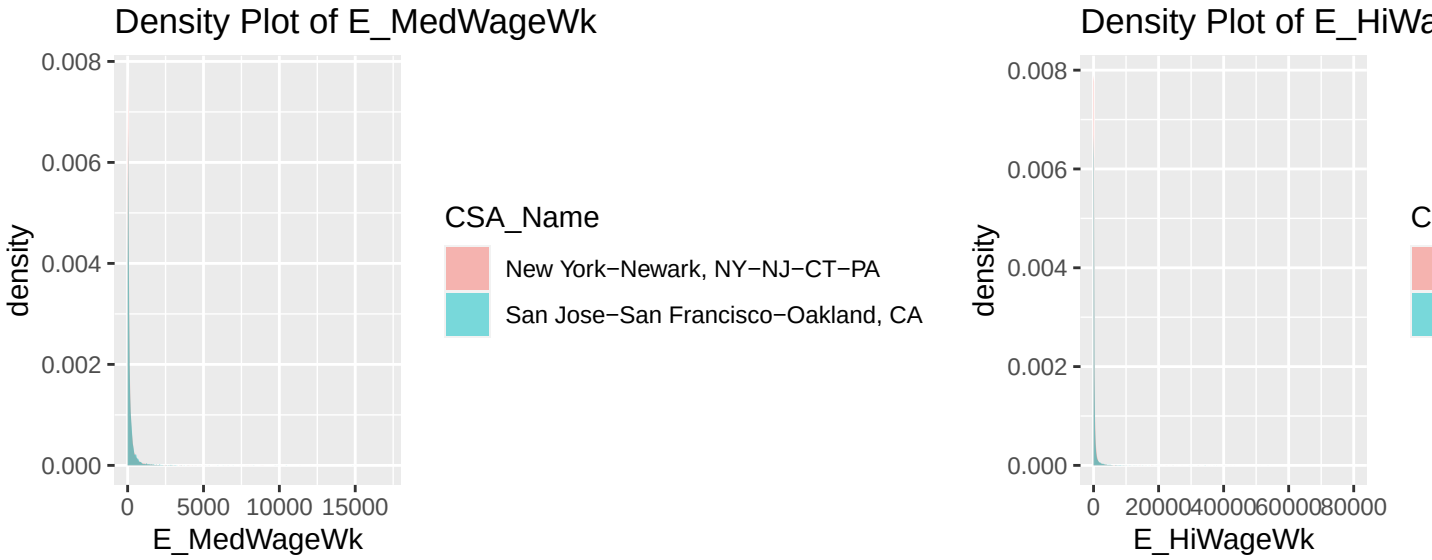

Density Plot of E8_Ent
density
0.04
0.03
0.02
0.01
0.00
0
2500
5000
7500
10000
E8_Ent

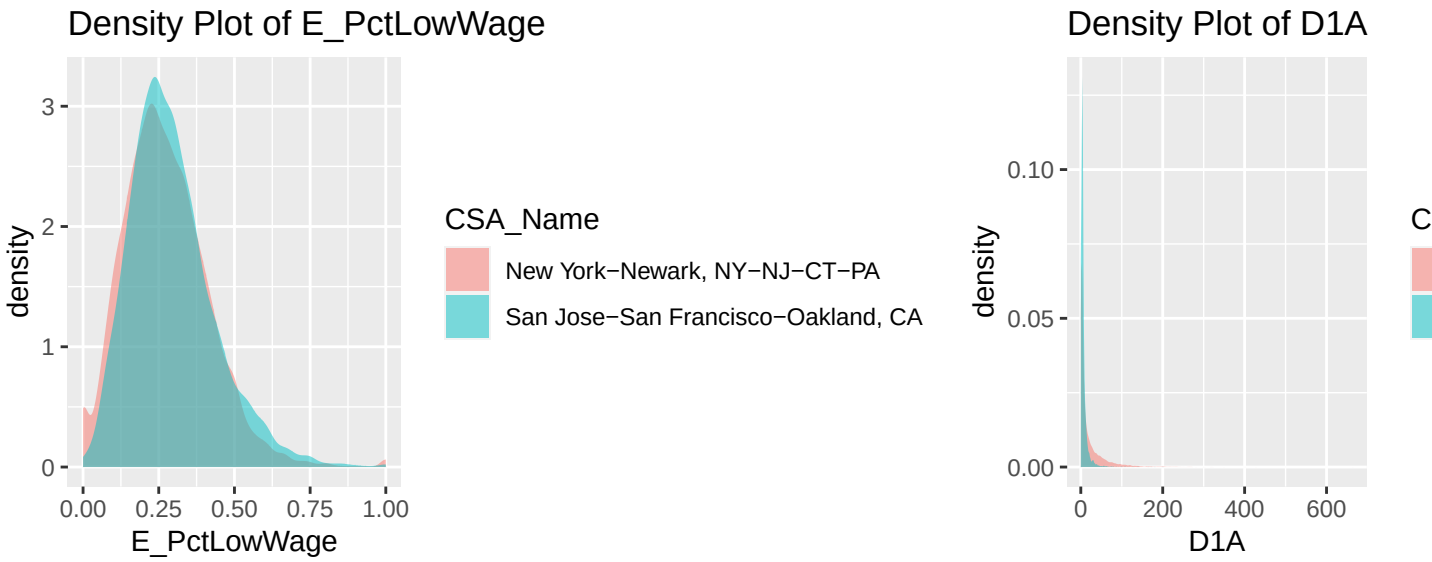

Density Plot of E8_Ed
density
0.075
0.050
0.025
0.000
0
10000
20000
E8_Ed
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA


Density Plot of E8_Hlth
density
0.03
0.02
0.01
0.00
0
10000
20000
30000
E8_Hlth

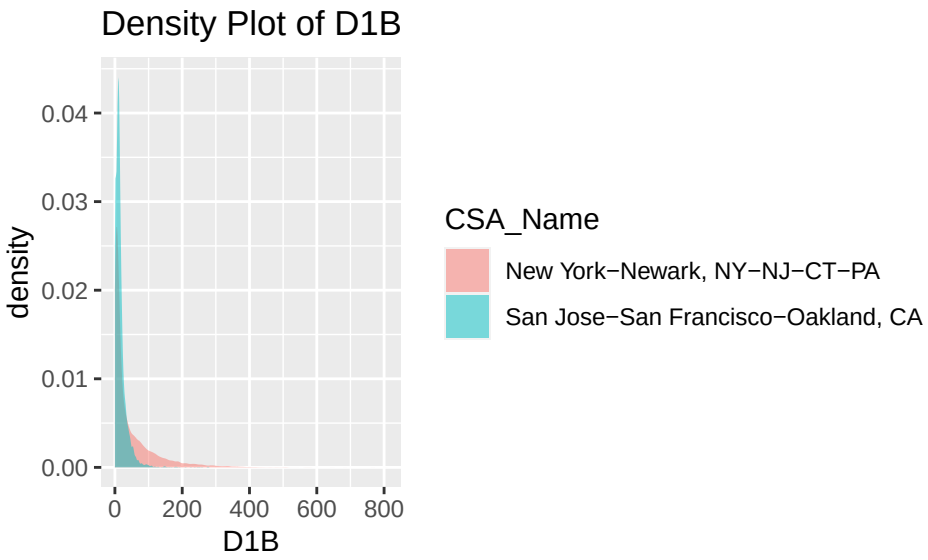

Density Plot of E8_Pub
density
0.0075
0.0050
0.0025
0.0000
0
10000
20000
30000
E8_Pub
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

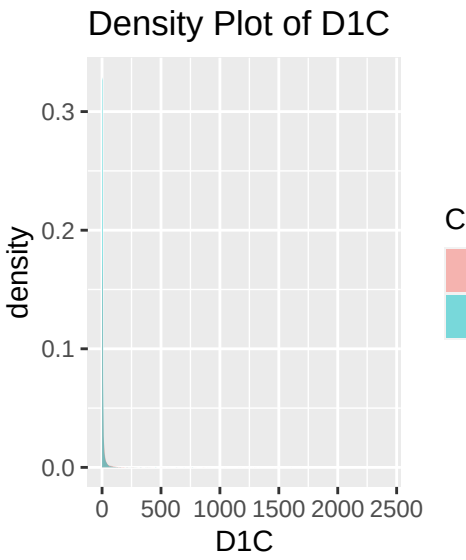

Density Plot of E_Low
density
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
0
5000
10000
15000
20000
E_LowWageWk

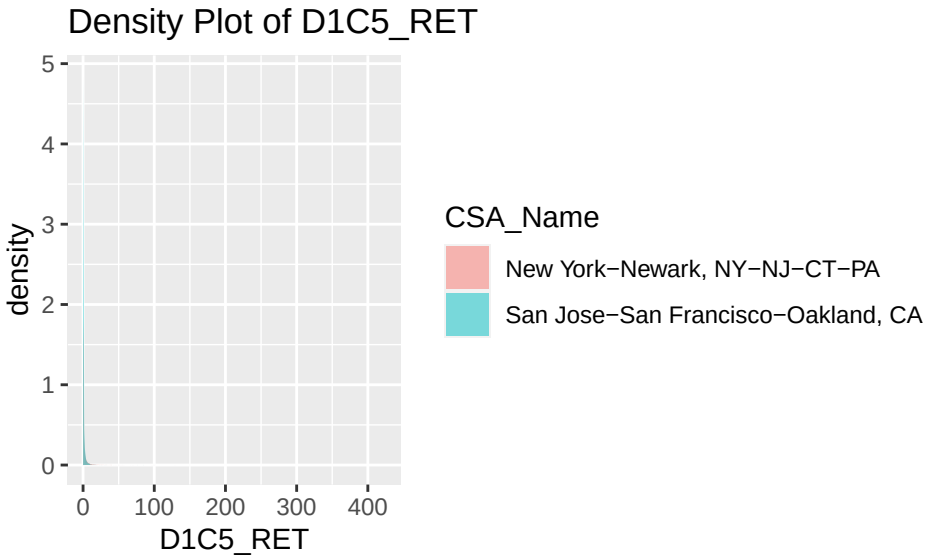


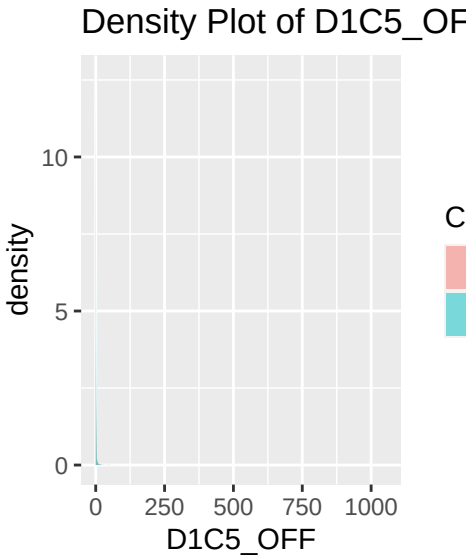


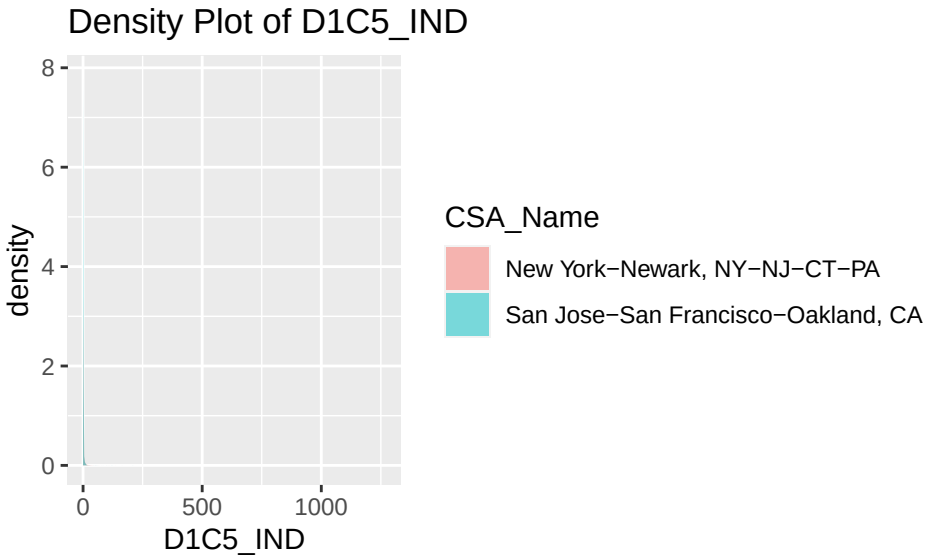

```
create_density_plots_subset(NewYork_SanFrancisco_data, 52, 76)
```

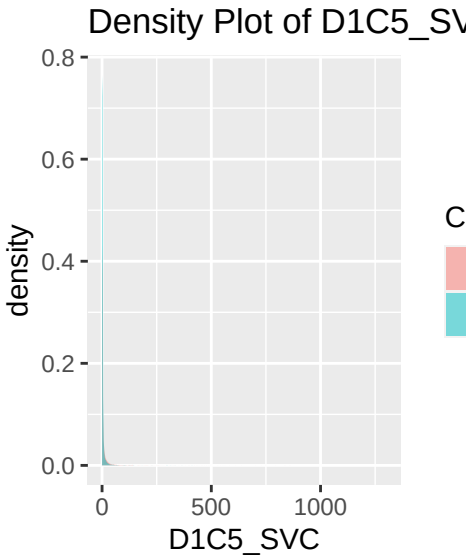

Density Plot of D1B
density
0.04
0.03
0.02
0.01
0.00
0
200
400
600
800
D1B
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

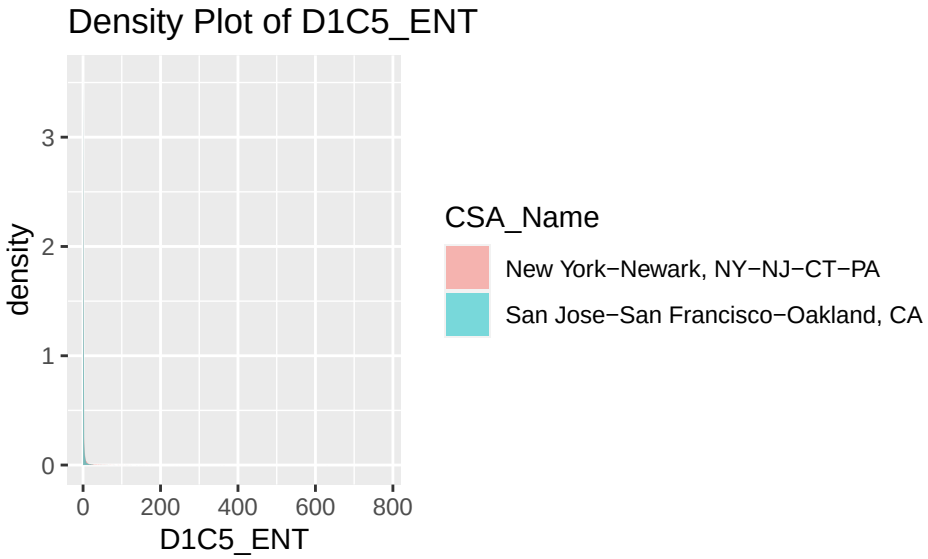

Density Plot of D1C
density
0.3
0.2
0.1
0.0
0
500
1000
1500
2000
2500
D1C

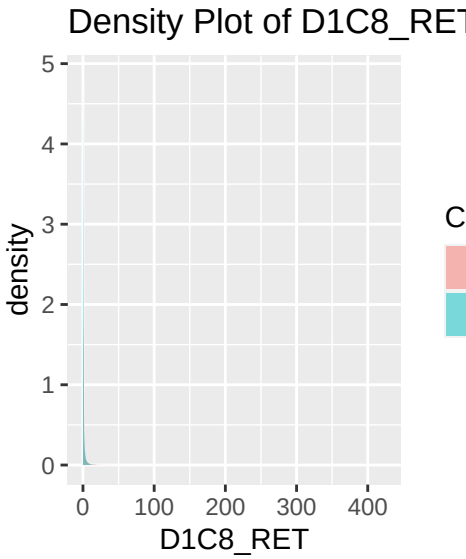

Density Plot of D1C5_RET
density
5
4
3
2
1
0
0
100
200
300
400
D1C5_RET
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

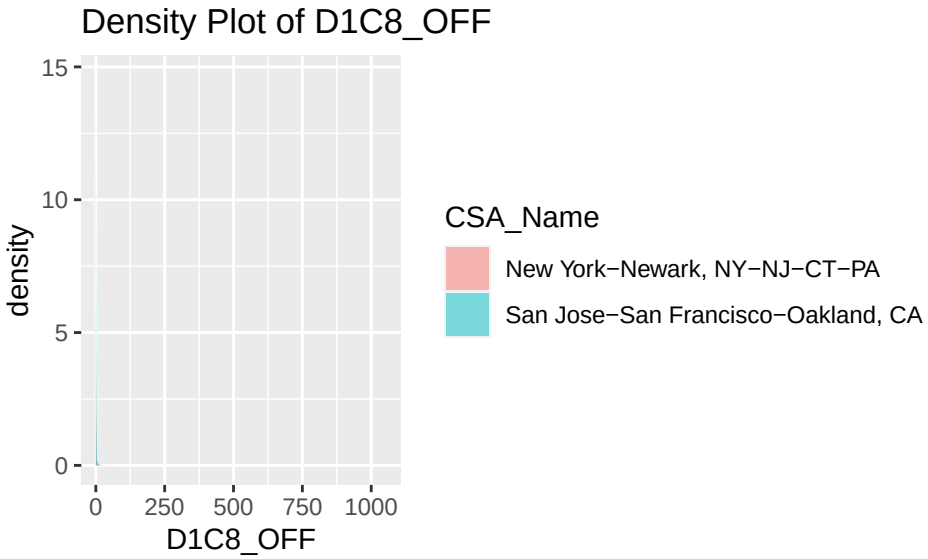

Density Plot of D1C5_OF
density
10
5
0
0
250
500
750
1000
D1C5_OFF

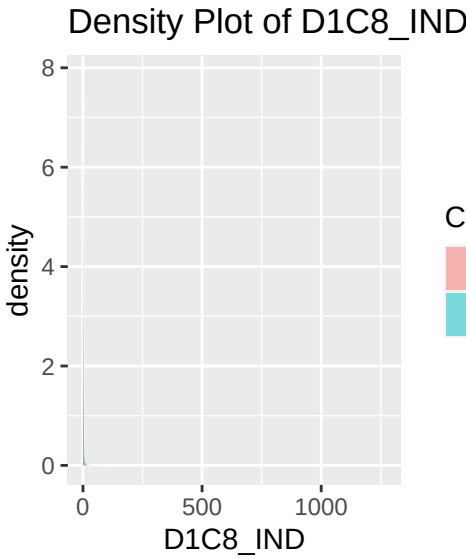

Density Plot of D1C5_IND
density
D1C5_IND
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

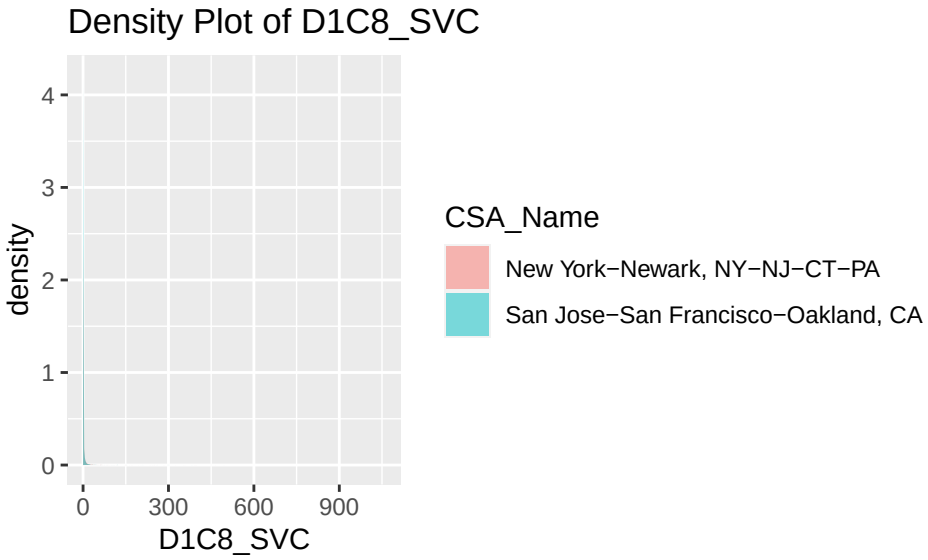

Density Plot of D1C5_SV
density
D1C5_SVC

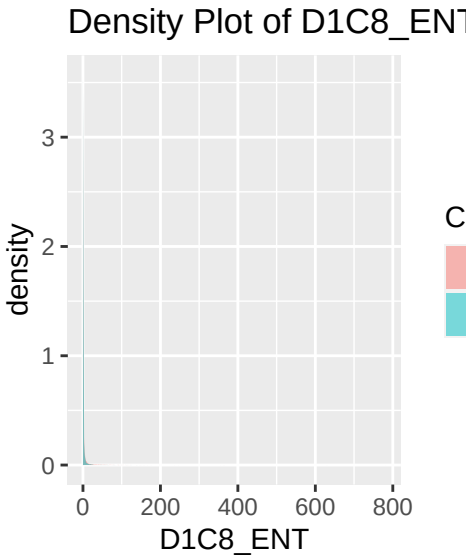

Density Plot of D1C5_ENT
density
D1C5_ENT
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

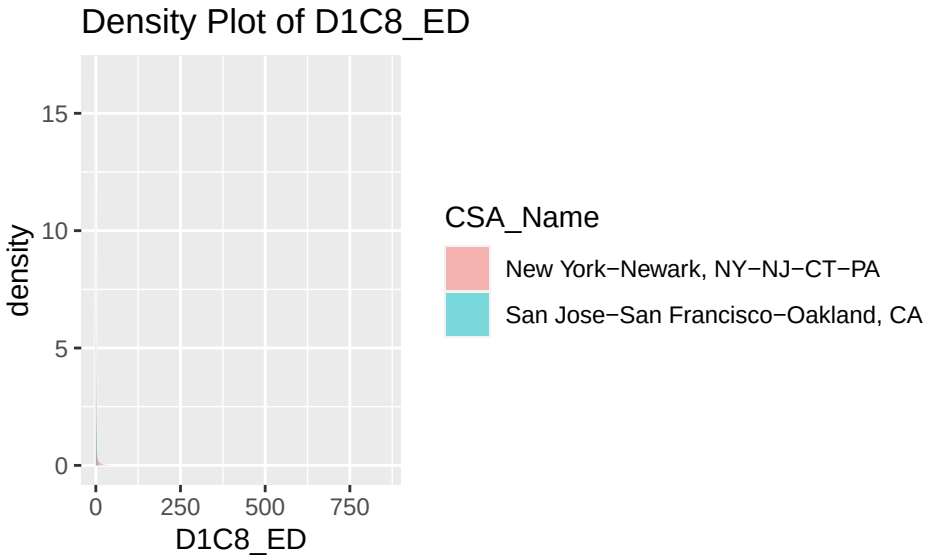

Density Plot of D1C8_RET
density
D1C8_RET

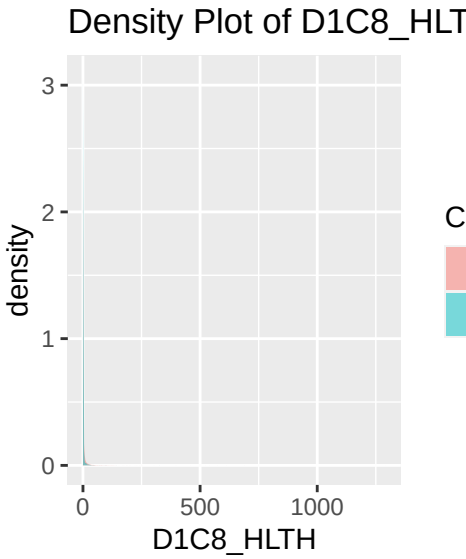

Density Plot of D1C8_OFF
density
D1C8_OFF
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

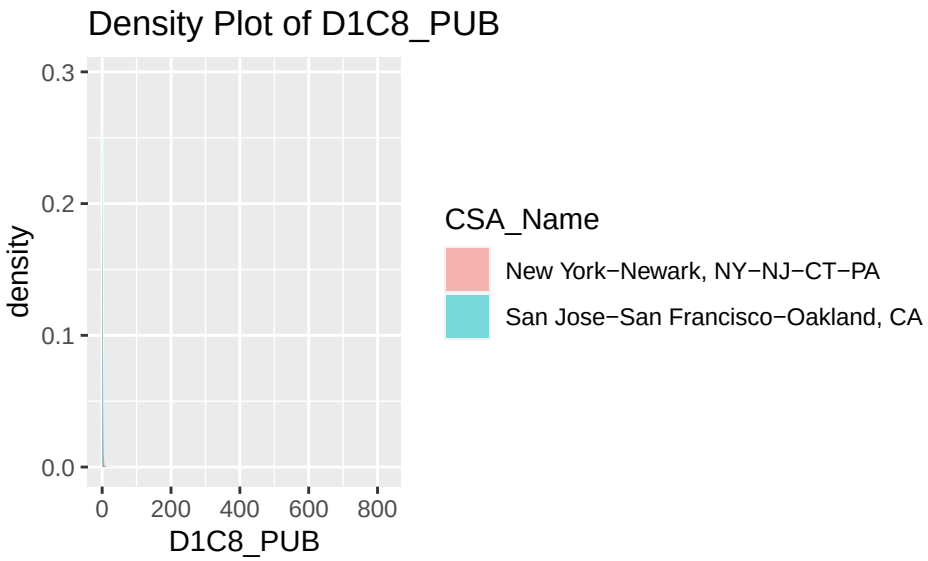

Density Plot of D1C8_IND
density
D1C8_IND

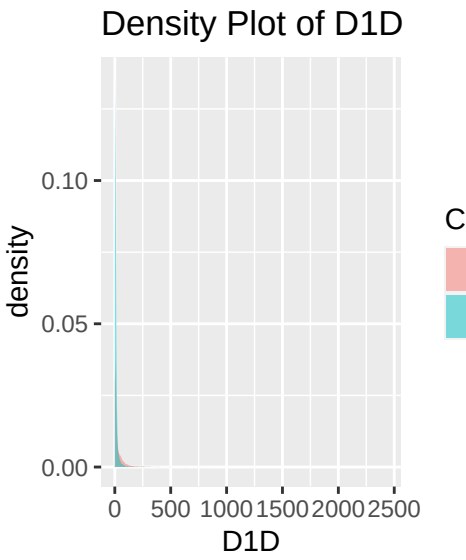

Density Plot of D1C8_SVC
density
D1C8_SVC
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

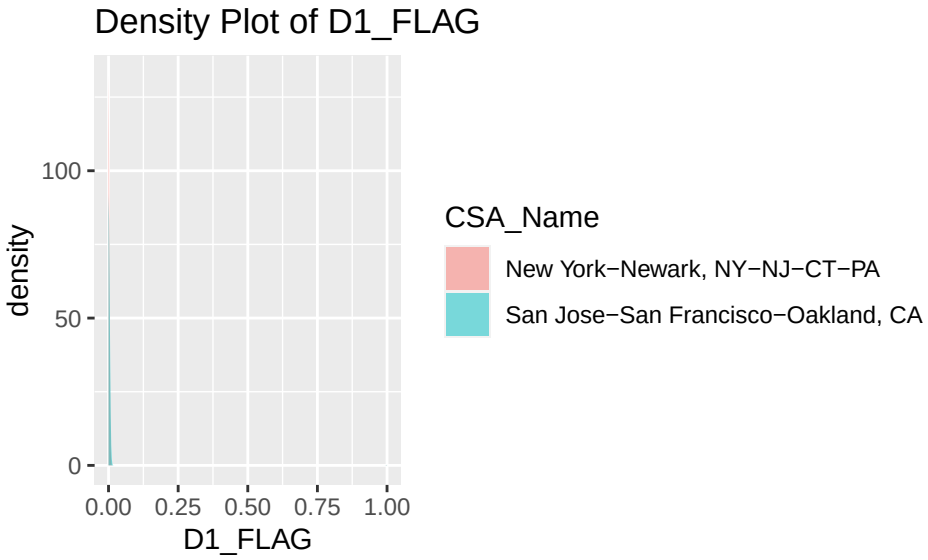

Density Plot of D1C8_ENT
density
D1C8_ENT

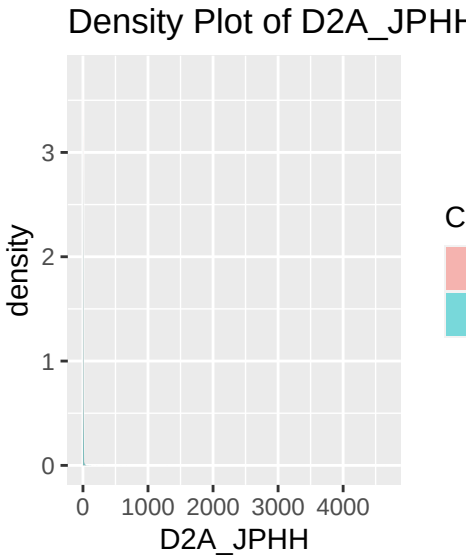

Density Plot of D1C8_ED
density
15
10
5
0
0
250
500
750
D1C8_ED
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

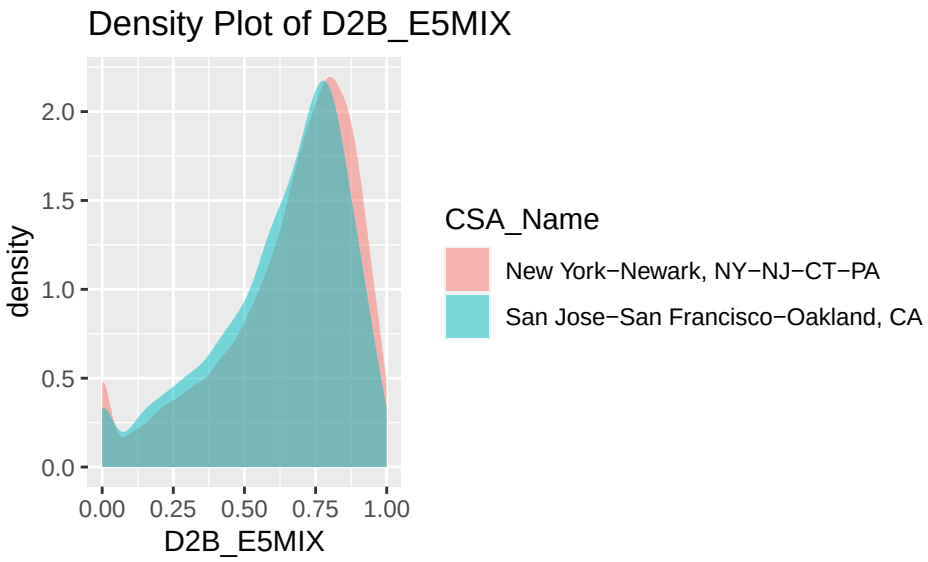

Density Plot of D1C8_HLT
density
3
2
1
0
0
500
1000
D1C8_HLTH
C

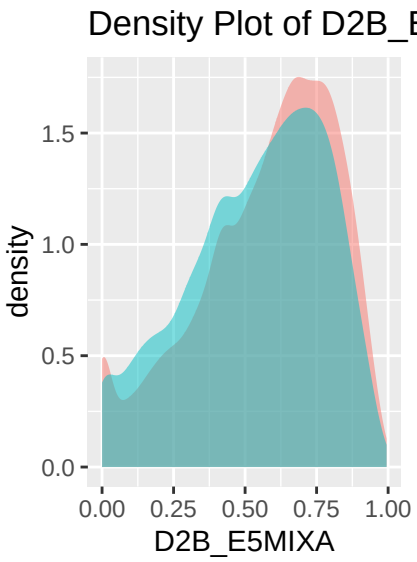

Density Plot of D1C8_PUB
density
0.3
0.2
0.1
0.0
0
200
400
600
800
D1C8_PUB
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

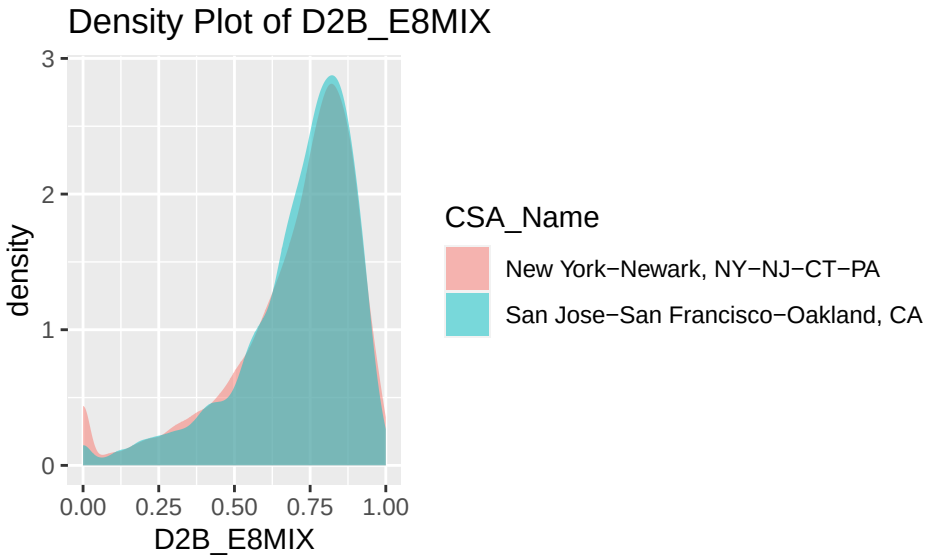

Density Plot of D1D
density
0.10
0.05
0.00
0
500
1000
1500
2000
2500
D1D
C

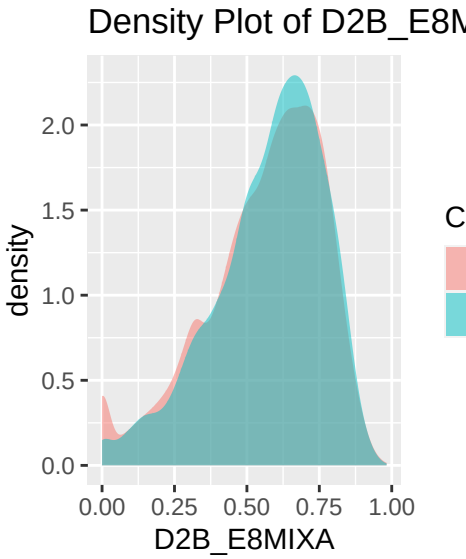

Density Plot of D1_FLAG
density
100
50
0
0.00
0.25
0.50
0.75
1.00
D1_FLAG
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

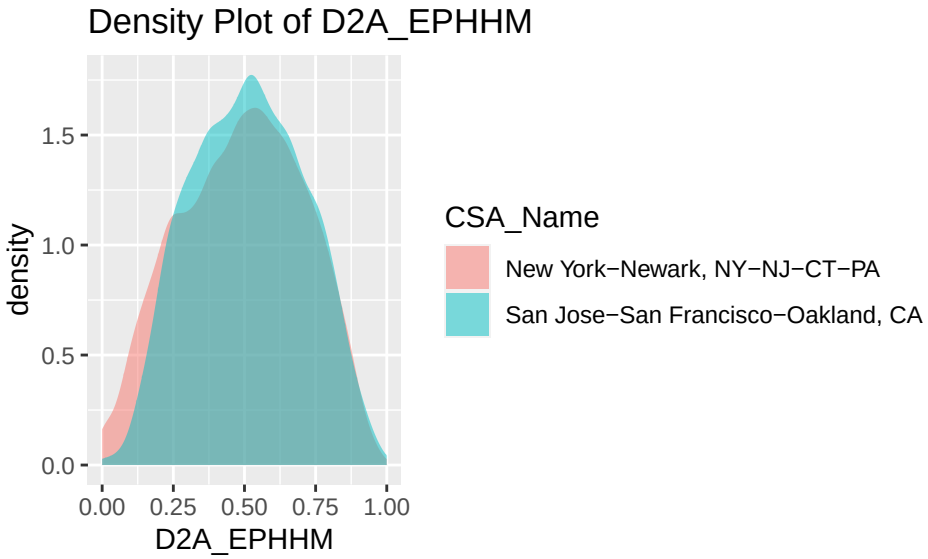

Density Plot of D2A_JPHH
density
3
2
1
0
0
1000
2000
3000
4000
D2A_JPHH

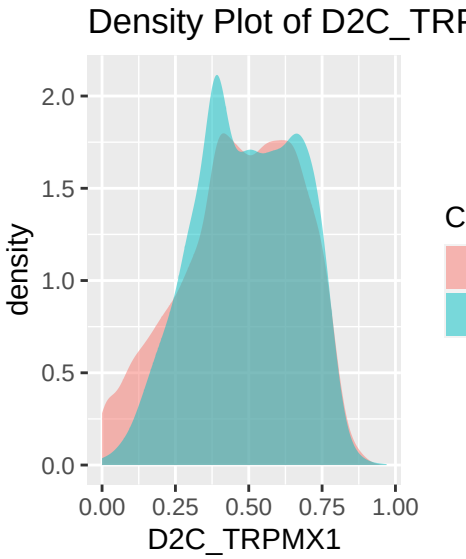

Density Plot of D2B_E5MIX
density
2.0
1.5
1.0
0.5
0.0
0.00
0.25
0.50
0.75
1.00
D2B_E5MIX
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

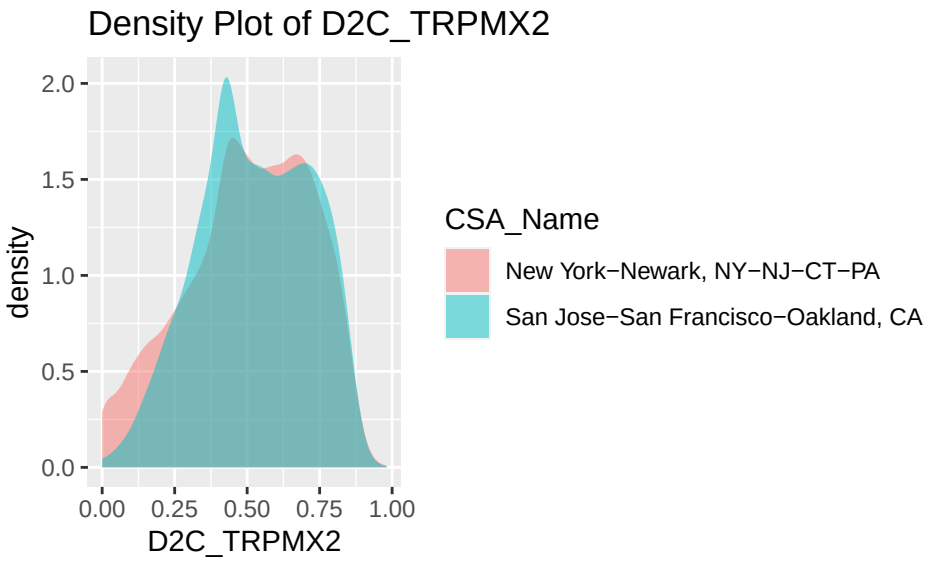

Density Plot of D2B_E5M
density
1.5
1.0
0.5
0.0
0.00
0.25
0.50
0.75
1.00
D2B_E5MIXA


Density Plot of D2B_E8MIX
density
D2B_E8MIX
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

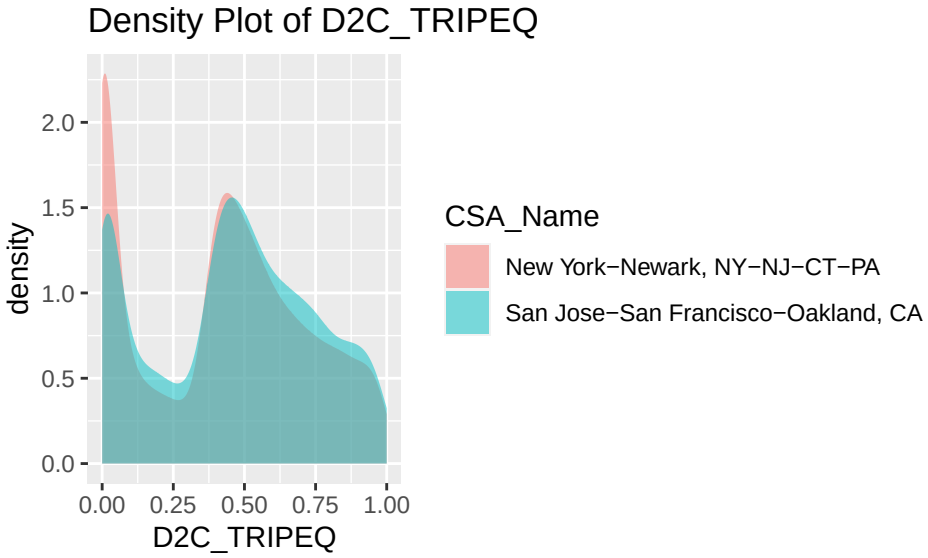

Density Plot of D2B_E8M
density
D2B_E8MIXA

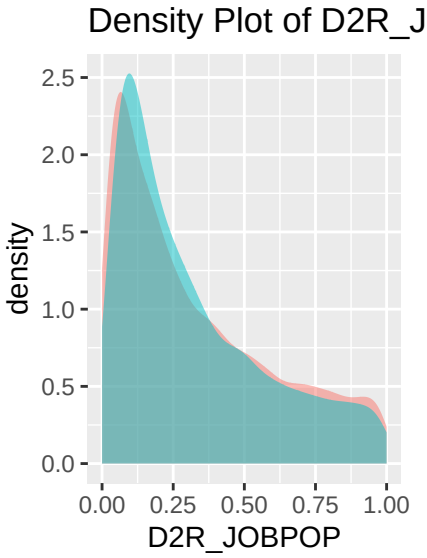

Density Plot of D2A_EPHHM
density
D2A_EPHHM
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

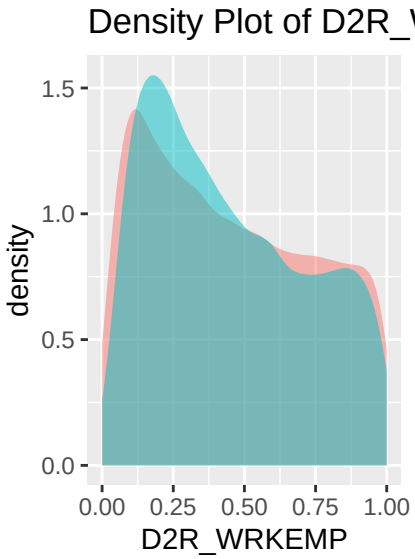

Density Plot of D2C_TRP
density
D2C_TRPMX1

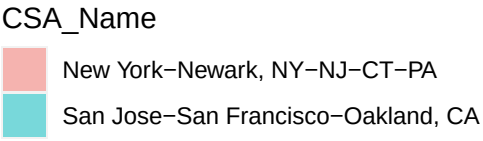


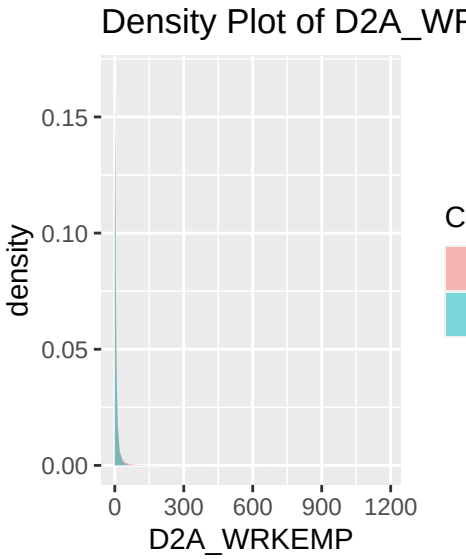

```
create_density_plots_subset(NewYork_SanFrancisco_data, 77, 118)
```

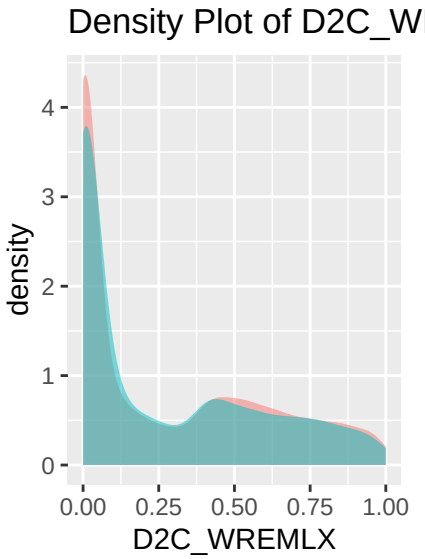


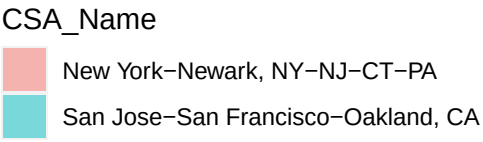


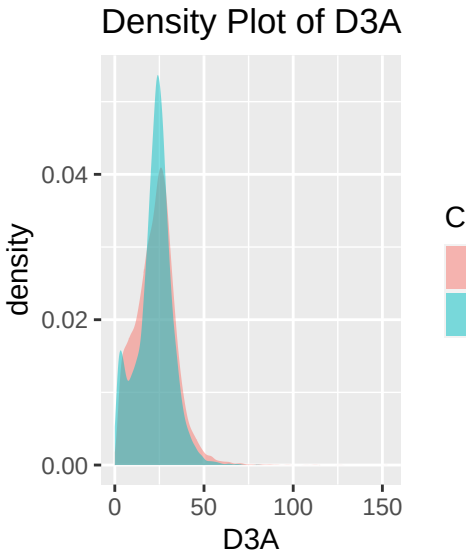

Density Plot of D2R_WRKEMP
density
1.5
1.0
0.5
0.0
0.00
0.25
0.50
0.75
1.00
D2R_WRKEMP

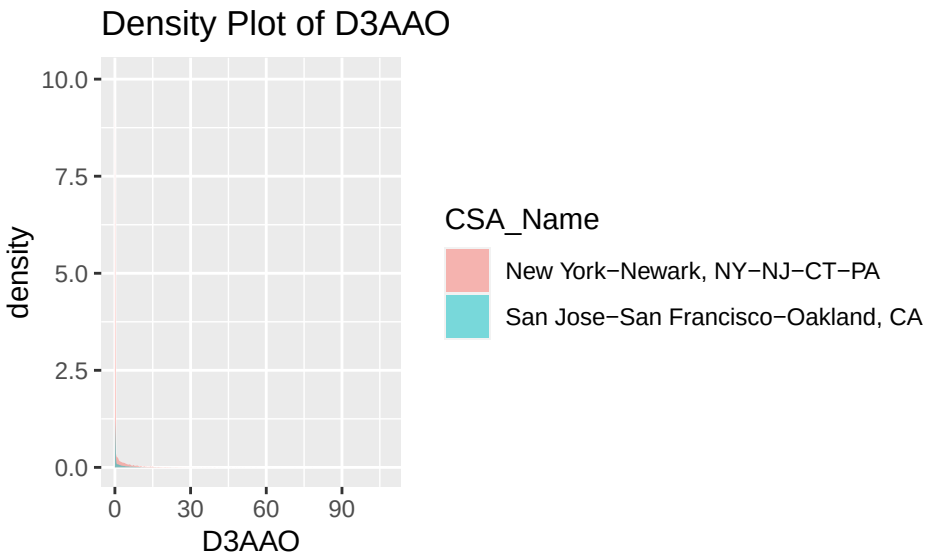

CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

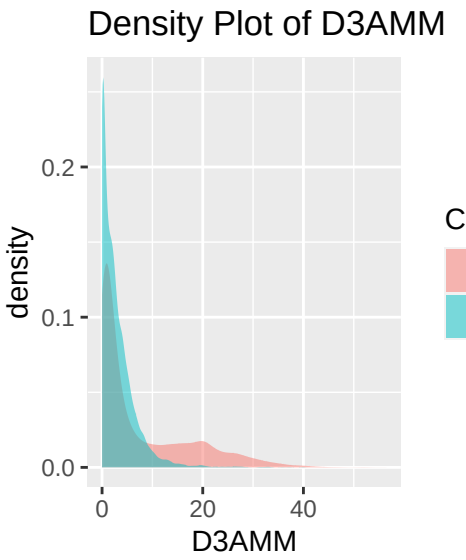

Density Plot of D2A_WR
density
0.15
0.10
0.05
0.00
0
300
600
900
1200
D2A_WRKEMP

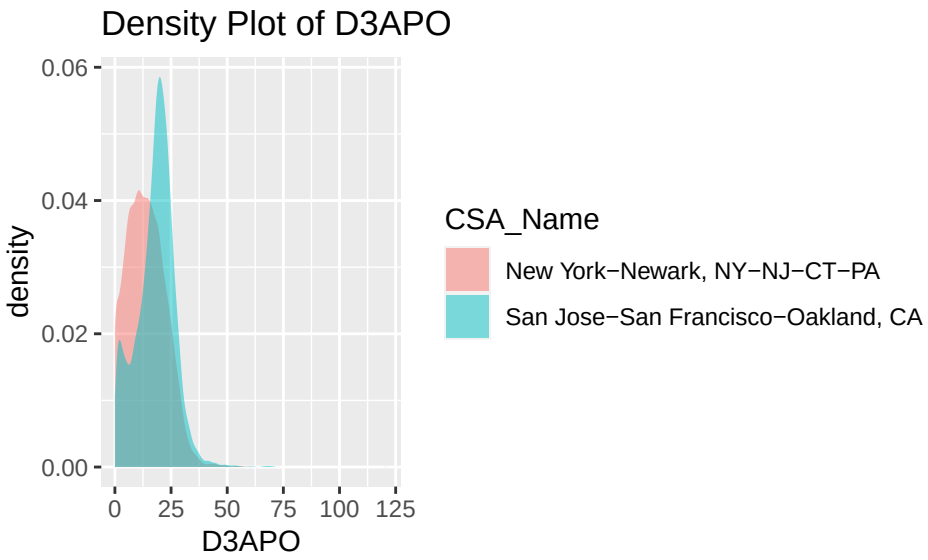

Density Plot of D2C_WREMLX
density
4
3
2
1
0
0.00
0.25
0.50
0.75
1.00
D2C_WREMLX

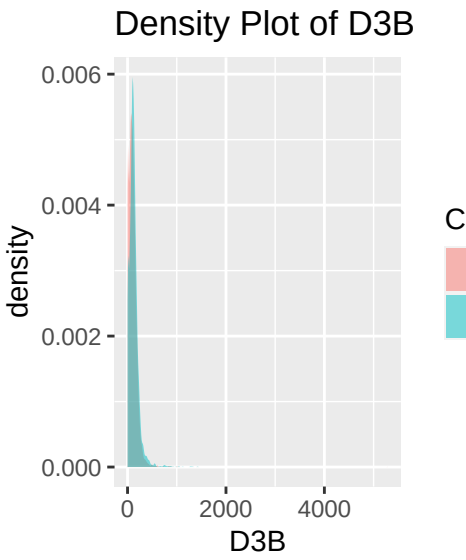

CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

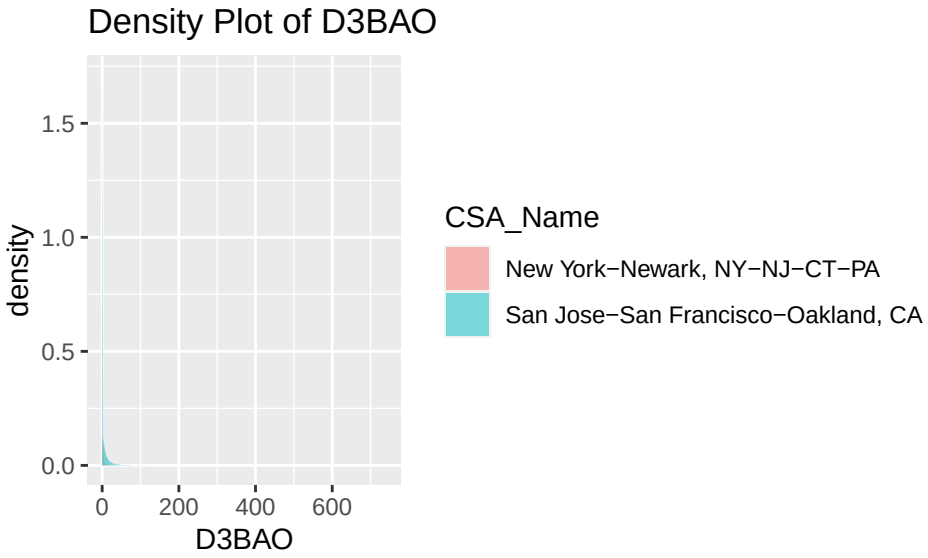

Density Plot of D3A
density
0.04
0.02
0.00
0
50
100
150
D3A

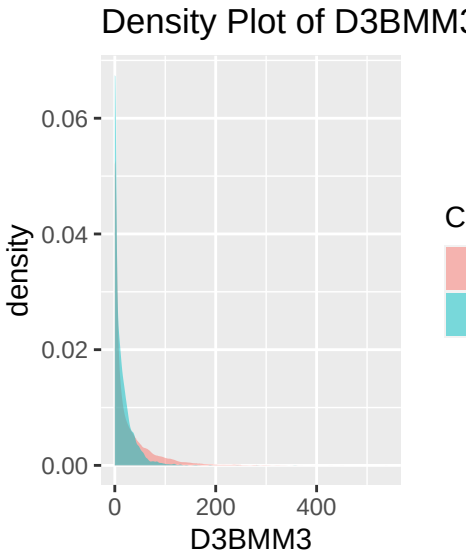

Density Plot of D3AAO
density
10.0
7.5
5.0
2.5
0.0
0
30
60
90
D3AAO
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

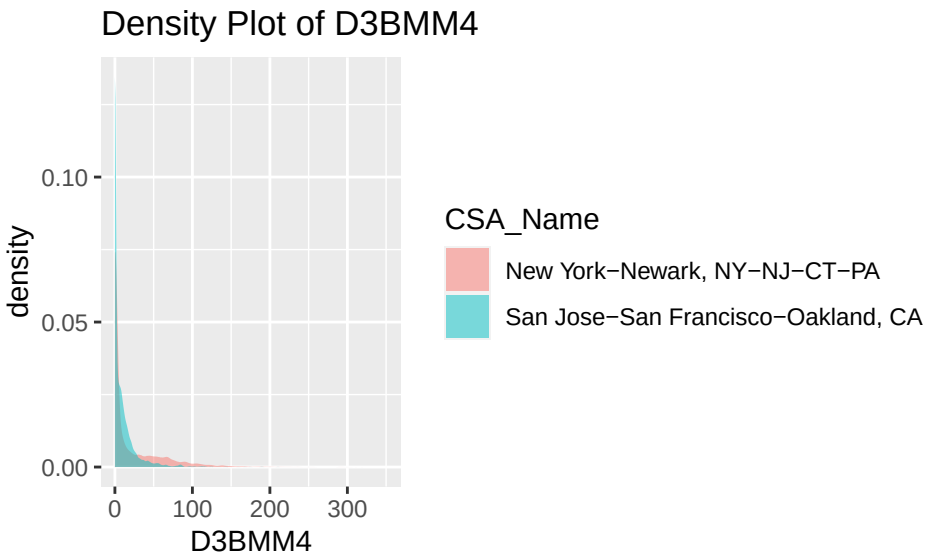

Density Plot of D3AMM
density
0.2
0.1
0.0
0
20
40
D3AMM
C

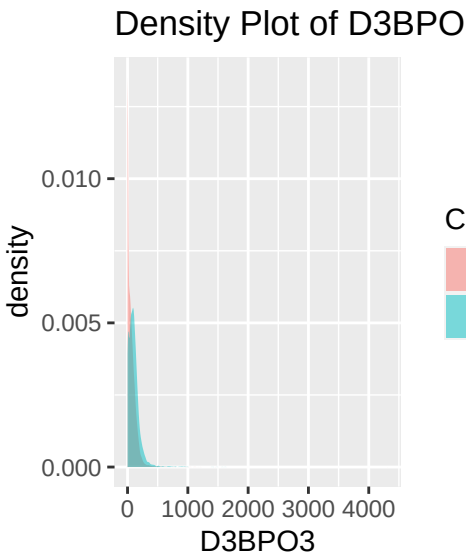

Density Plot of D3APO
density
0.06
0.04
0.02
0.00
0
25
50
75
100
125
D3APO
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

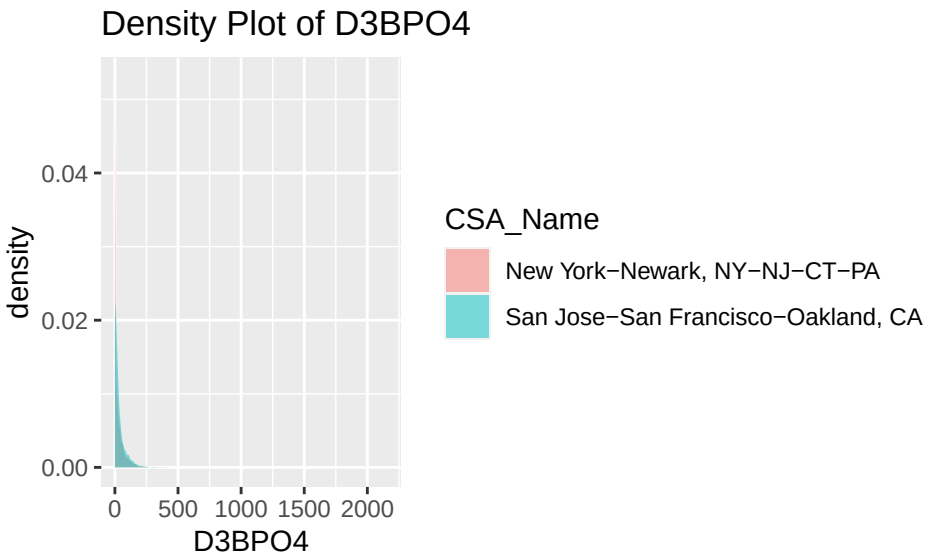

Density Plot of D3B
density
0.006
0.004
0.002
0.000
0
2000
4000
D3B
C

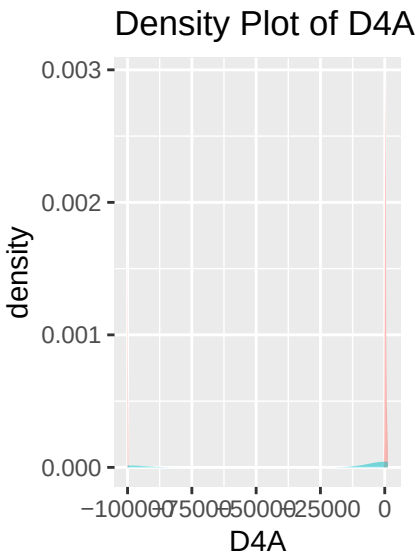

Density Plot of D3BAO
density
1.5
1.0
0.5
0.0
0
200
400
600
D3BAO
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

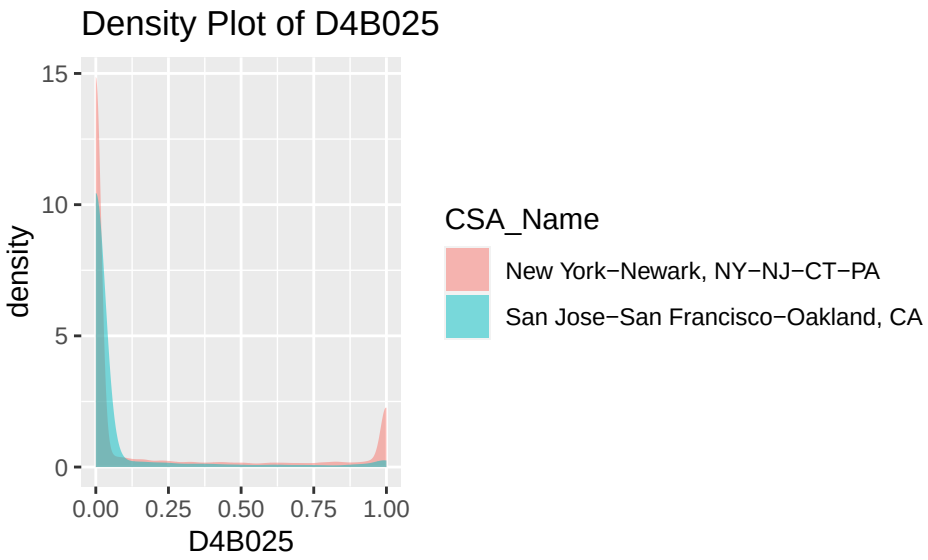

Density Plot of D3BMM3
density
0.06
0.04
0.02
0.00
0
200
400
D3BMM3

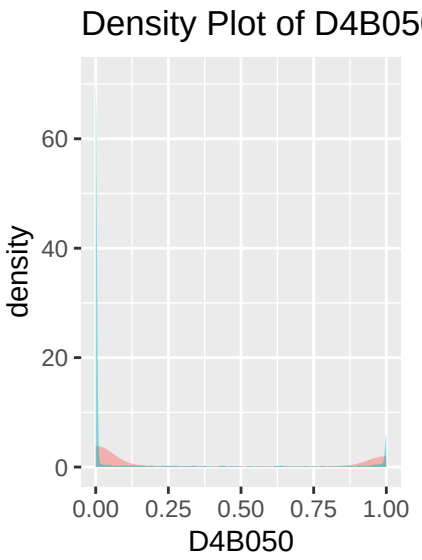

Density Plot of D3BMM4
density
0.10
0.05
0.00
0
100
200
300
D3BMM4
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

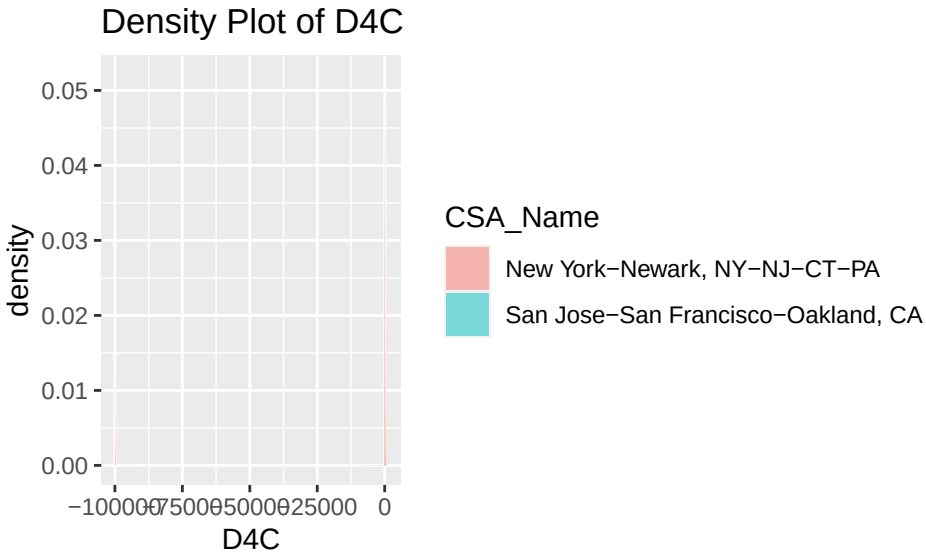

Density Plot of D3BPO
density
0.010
0.005
0.000
0
1000
2000
3000
4000
D3BPO3

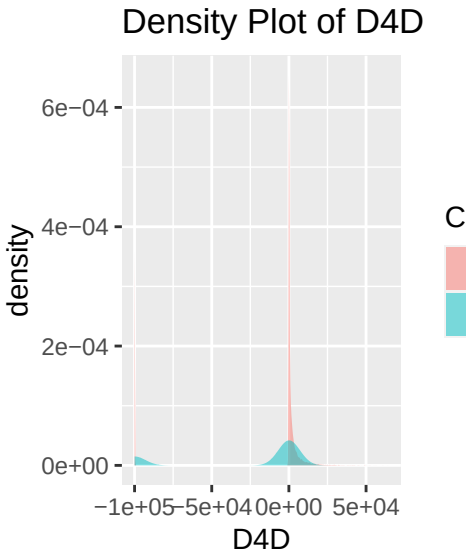

Density Plot of D3BPO4
density
0.04
0.02
0.00
0
500
1000
1500
2000
D3BPO4
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

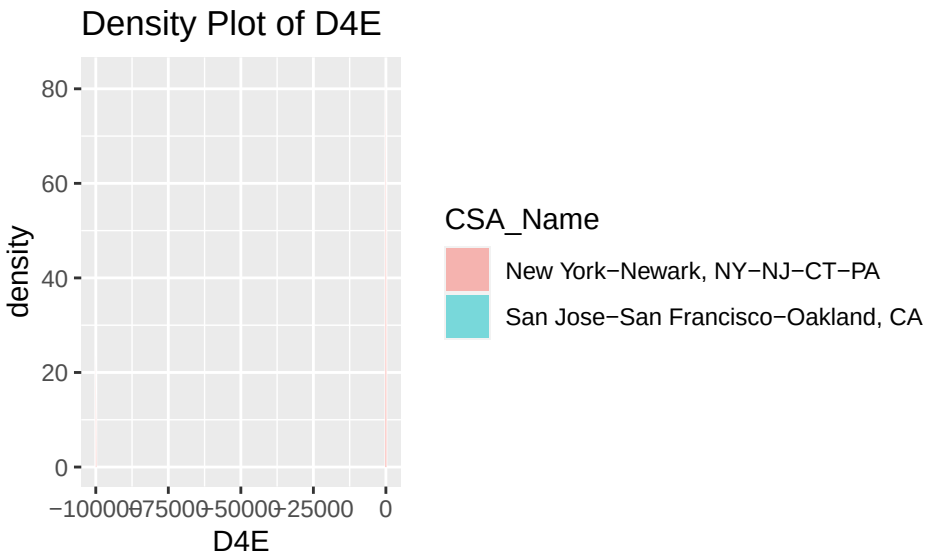

Density Plot of D4A
density
0.003
0.002
0.001
0.000
−100000
−75000
−50000
−25000
0
D4A

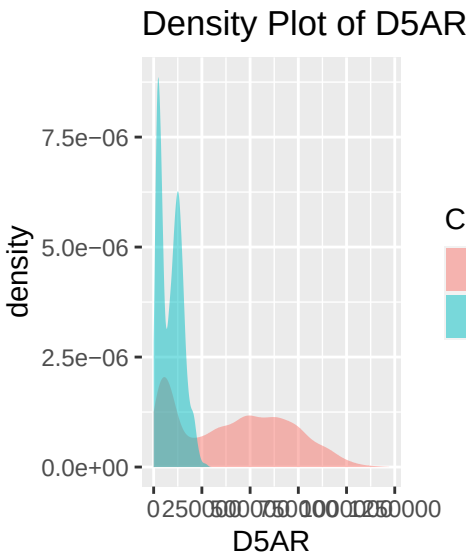

Density Plot of D4B025
density
15
10
5
0
0.00
0.25
0.50
0.75
1.00
D4B025
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

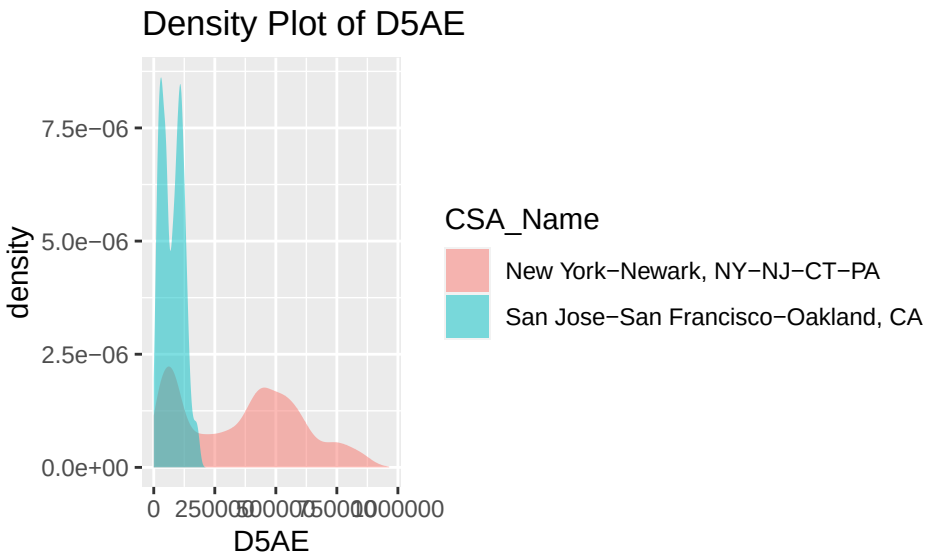

Density Plot of D4B050
density
60
40
20
0
0.00
0.25
0.50
0.75
1.00
D4B050

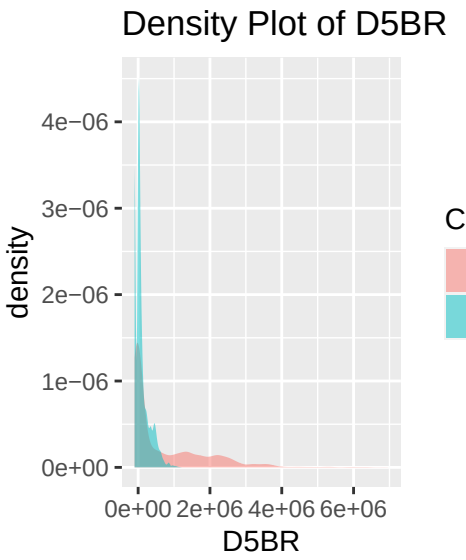

Density Plot of D4C
density
0.05
0.04
0.03
0.02
0.01
0.00
−100000
−75000
−50000
−25000
0
D4C
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

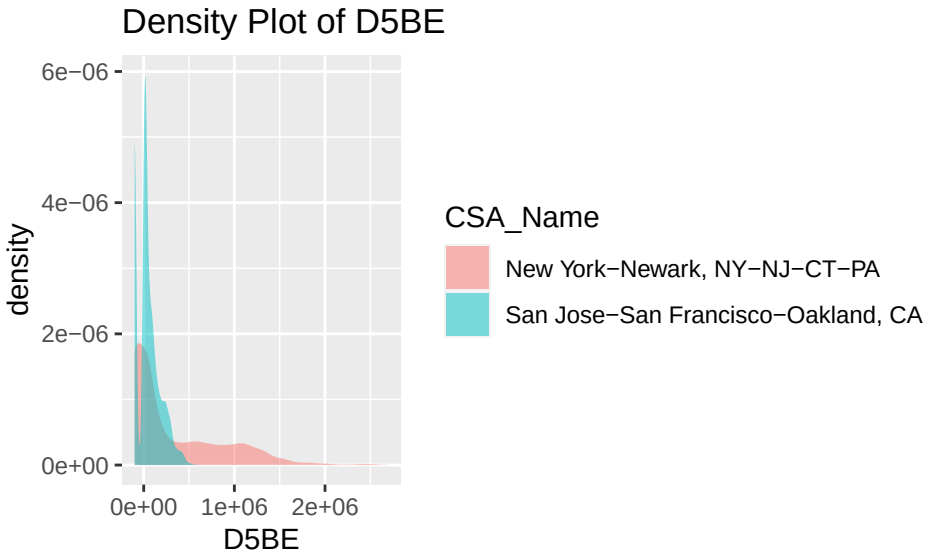

Density Plot of D4D
density
6e−04
4e−04
2e−04
0e+00
−1e+05
−5e+04
0e+00
5e+04
D4D

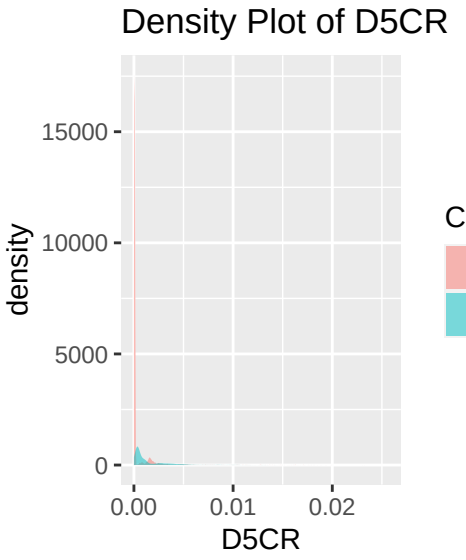

Density Plot of D4E
density
80
60
40
20
0
−100000
−75000
−50000
−25000
0
D4E
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

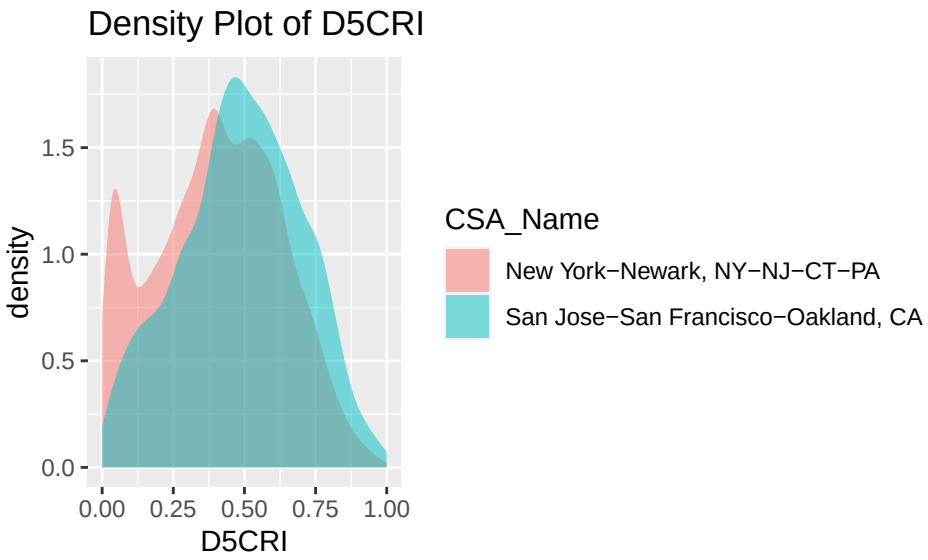

Density Plot of D5AR
density
7.5e−06
5.0e−06
2.5e−06
0.0e+00
D5AR

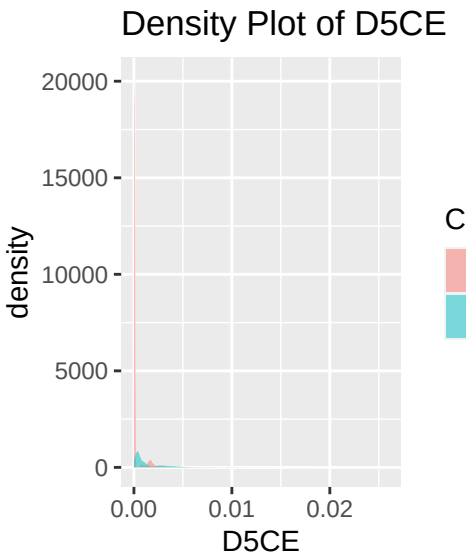

Density Plot of D5AE
density
7.5e−06
5.0e−06
2.5e−06
0.0e+00
0
D5AE
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

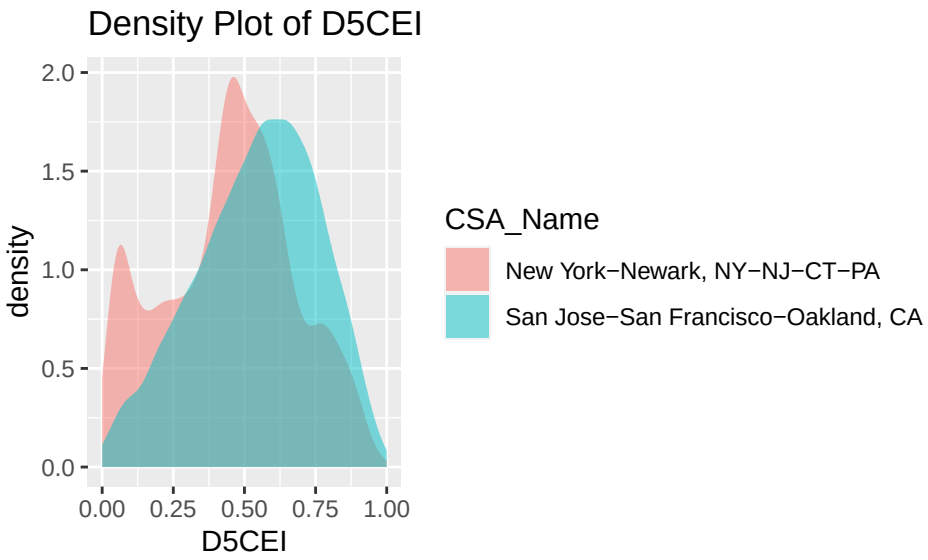

Density Plot of D5BR
density
4e−06
3e−06
2e−06
1e−06
0e+00
0e+00
2e+06
4e+06
6e+06
D5BR

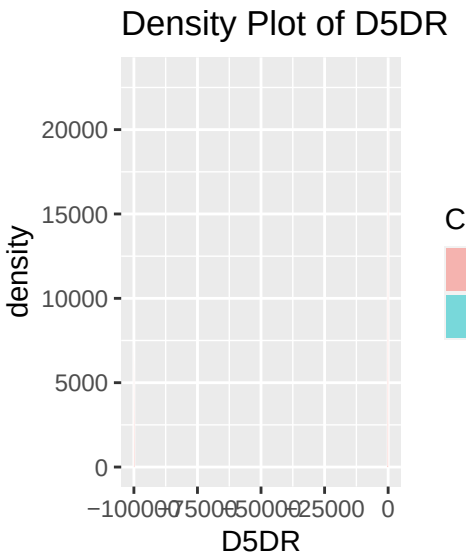

Density Plot of D5BE
density
6e−06
4e−06
2e−06
0e+00
0e+00
1e+06
2e+06
D5BE
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

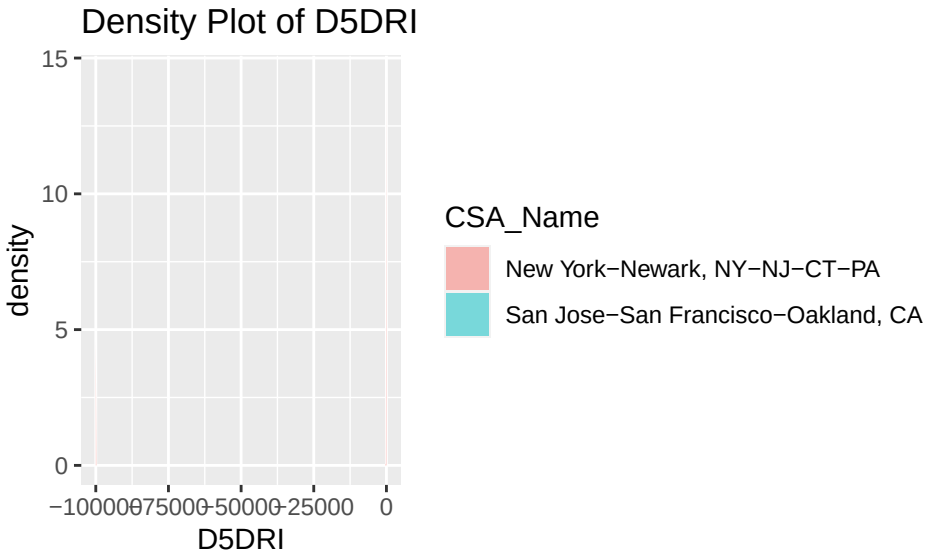

Density Plot of D5CR
density
15000
10000
5000
0
0.00
0.01
0.02
D5CR

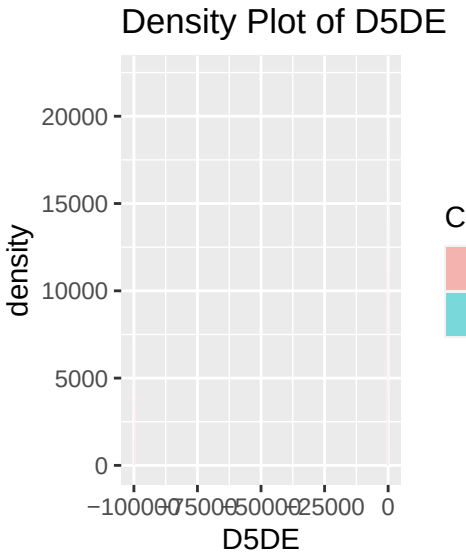

Density Plot of D5CRI
density
D5CRI
0.0
0.5
1.0
1.5
0.00
0.25
0.50
0.75
1.00
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

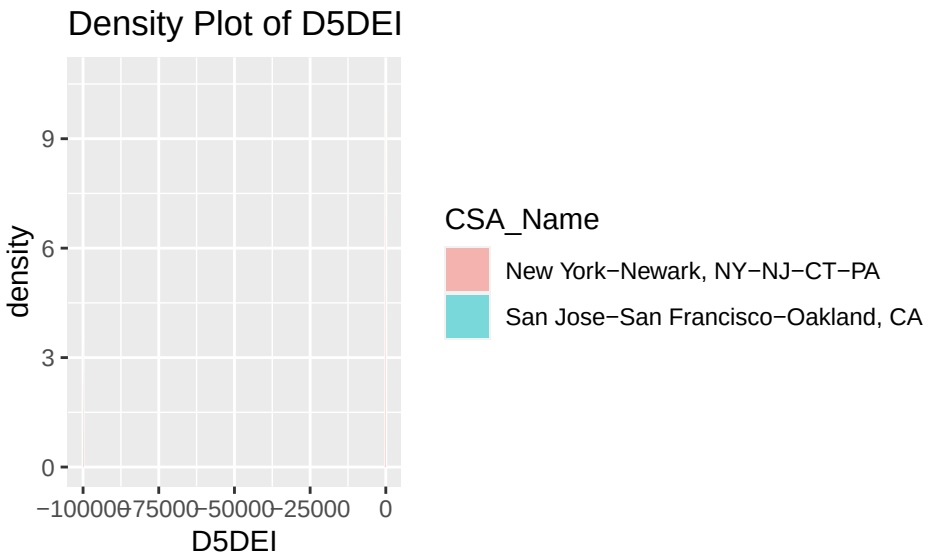

Density Plot of D5CE
density
D5CE
0
5000
10000
15000
20000
0.00
0.01
0.02

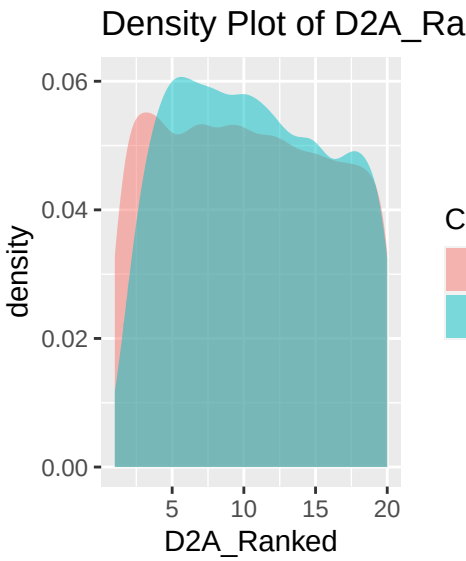

Density Plot of D5CEI
density
D5CEI
0.0
0.5
1.0
1.5
2.0
0.00
0.25
0.50
0.75
1.00
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

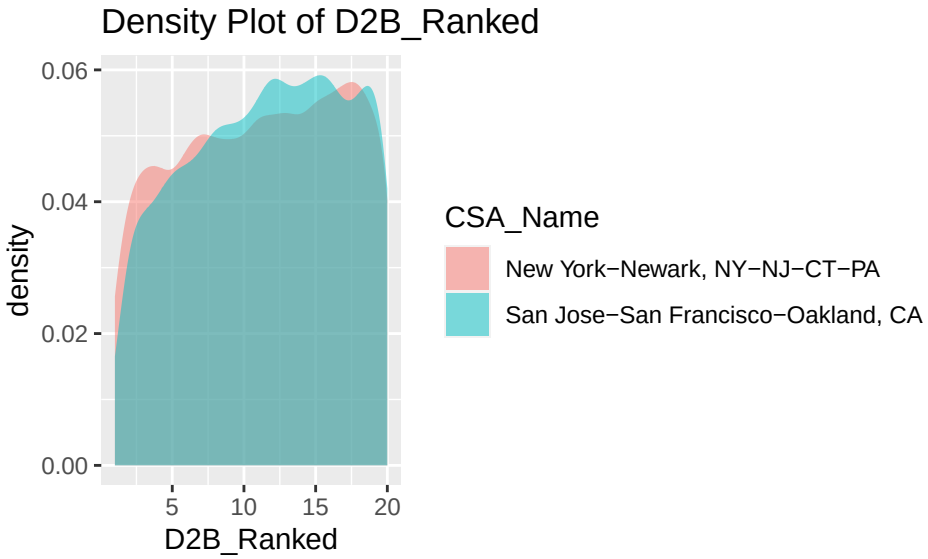

Density Plot of D5DR
density
D5DR
0
5000
10000
15000
20000
0

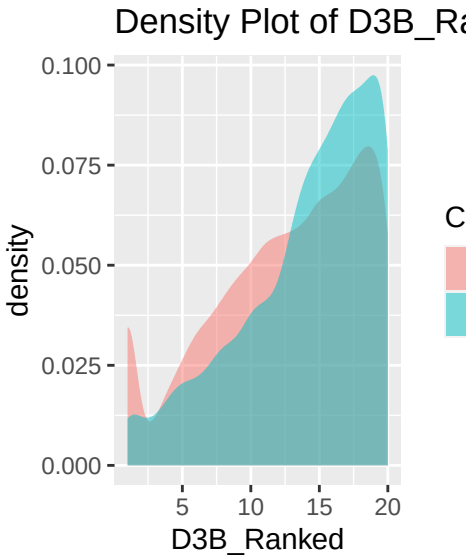

Density Plot of D5DRI
density
D5DRI
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

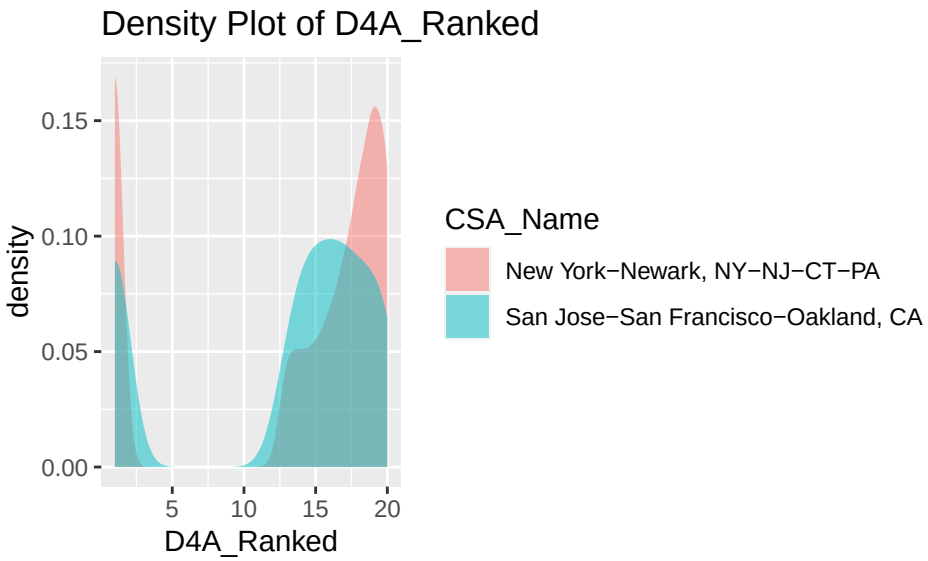

Density Plot of D5DE
density
D5DE

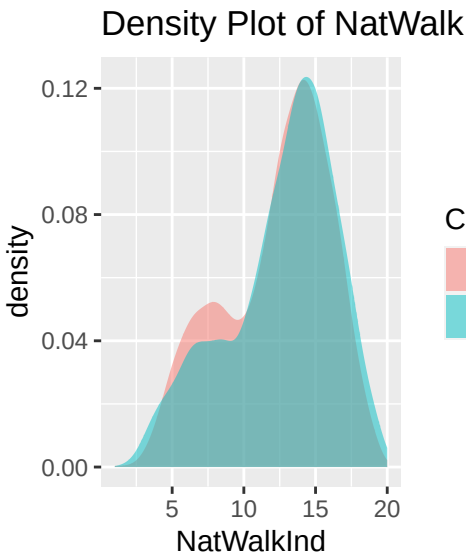

Density Plot of D5DEI
density
D5DEI
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

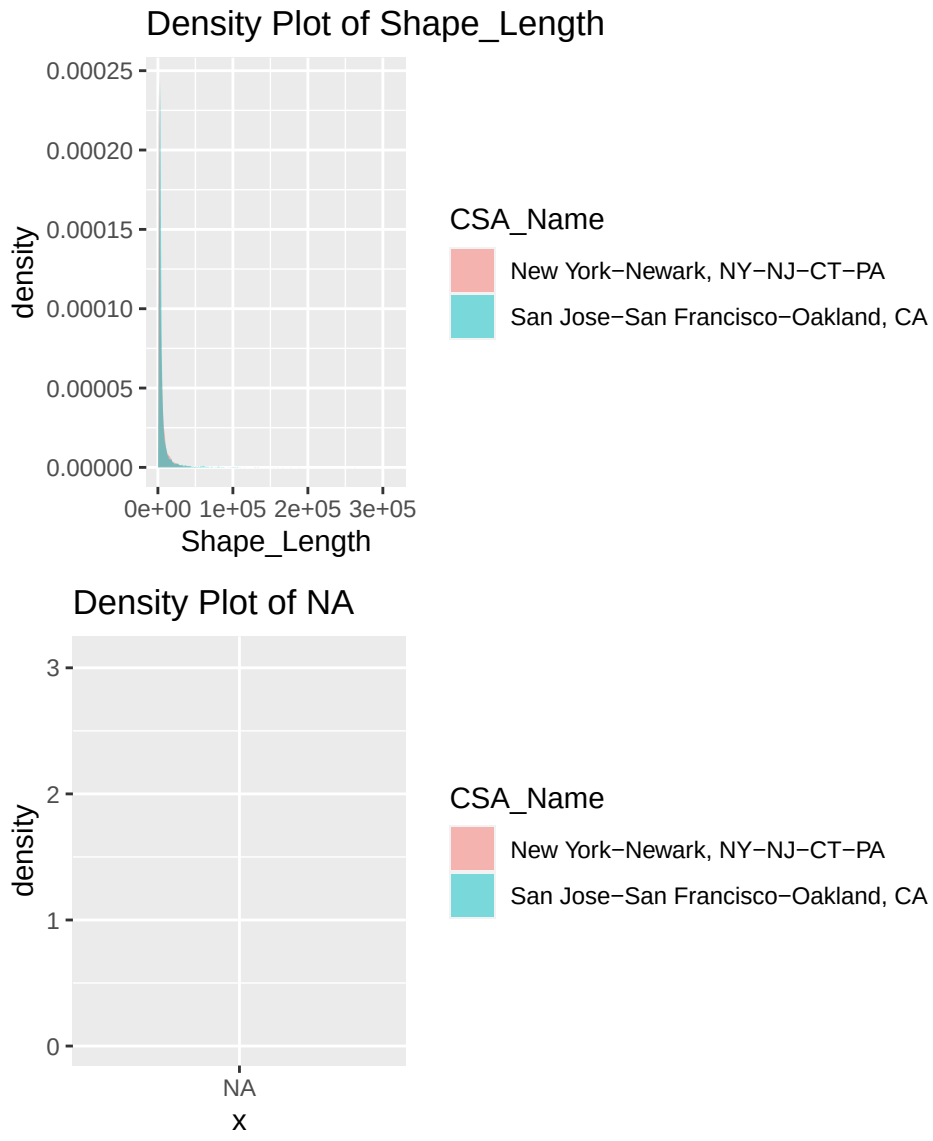

Density Plot of D2A_Ra
density
D2A_Ranked


Density Plot of D2B_Ranked
density
0.06
0.04
0.02
0.00
5
10
15
20
D2B_Ranked
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA


Density Plot of D3B_Ra
density
0.100
0.075
0.050
0.025
0.000
5
10
15
20
D3B_Ranked
C


Density Plot of D4A_Ranked
density
0.15
0.10
0.05
0.00
5
10
15
20
D4A_Ranked
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA


Density Plot of NatWalk
density
0.12
0.08
0.04
0.00
5
10
15
20
NatWalkInd
C

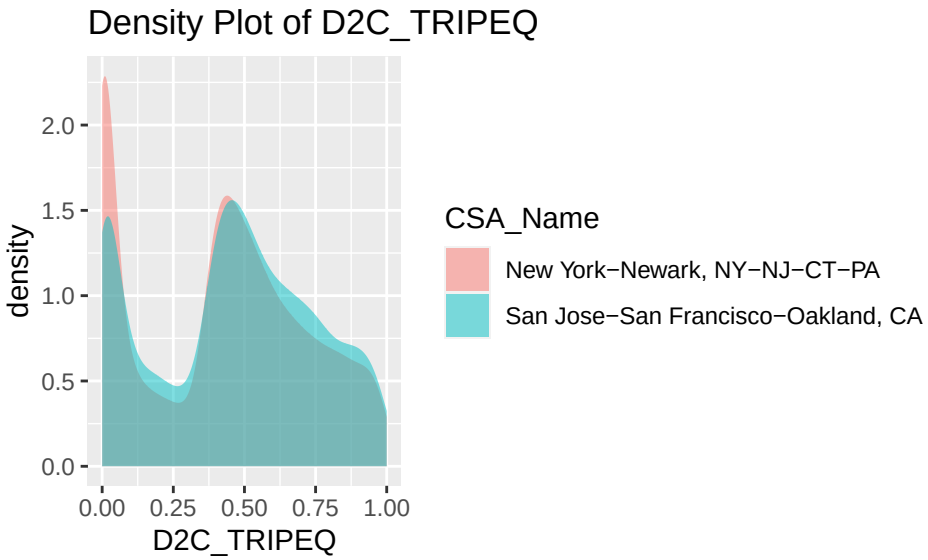


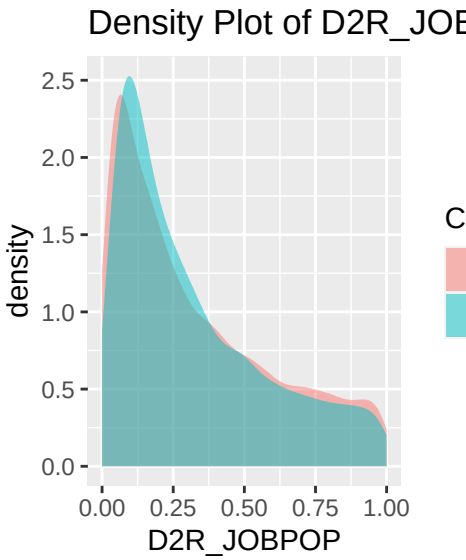
For columns 11 to 20:

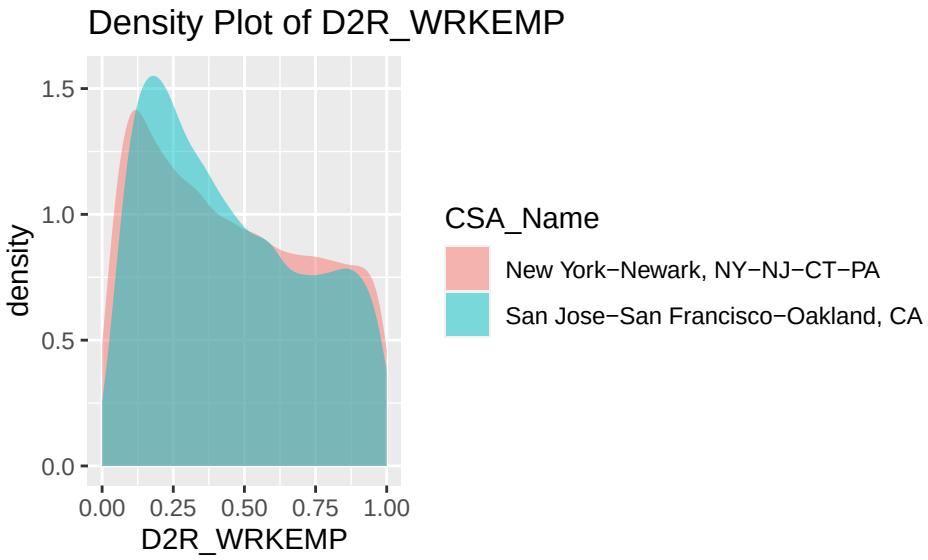
#Selecting Variables (there is way too many of them to run ggpairs on all of them)

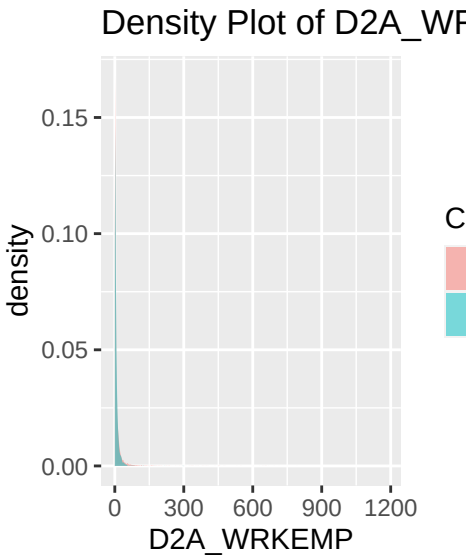
```
NewYork_SanFrancisco_data <- NewYork_SanFrancisco_data %>%
  mutate(CSA_Name_ABBREV = gsub("New York-Newark, NY-NJ-CT-PA", "NY", CSA_Name)) %>%
  mutate(CSA_Name_ABBREV = gsub("San Jose-San Francisco-Oakland, CA", "SF", CSA_Name))

# Remove leading and trailing whitespace
NewYork_SanFrancisco_data$CSA_Name_ABBREV <- trimws(NewYork_SanFrancisco_data$CSA_Name_ABBREV)

# Replace "New York-Newark, NY-NJ-CT-PA" with "NY"
NewYork_SanFrancisco_data$CSA_Name_ABBREV <- gsub("New York-Newark, NY-NJ-CT-PA", "NY", NewYork_SanFranc

# Replace "San Jose-San Francisco-Oakland, CA" with "SF"
```

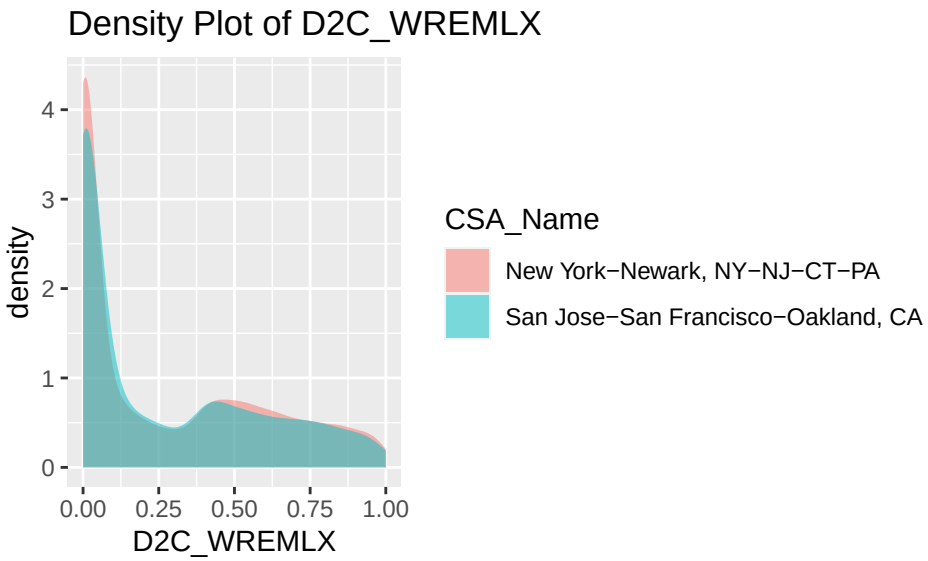
```
NewYork_SanFrancisco_data$CSA_Name_ABBREV <- gsub("San Jose-San Francisco-Oakland, CA", "SF", NewYork_Sa


NY_SF_focus <- NewYork_SanFrancisco_data %>%
  select(NatWalkInd, CSA_Name_ABBREV, Ac_Unpr, P_WrkAge, P_WrkAge, Pct_AO0, Pct_AO1, Pct_AO2p, R_LowWage

create_density_plots_subset(NewYork_SanFrancisco_data, 77, 118)
```

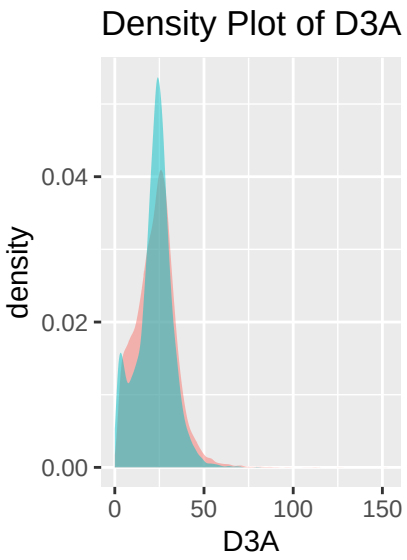


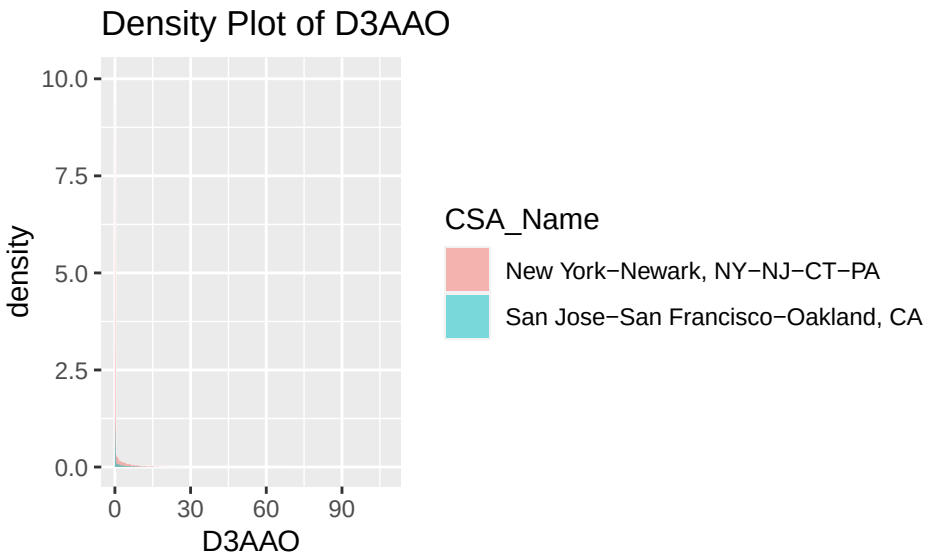


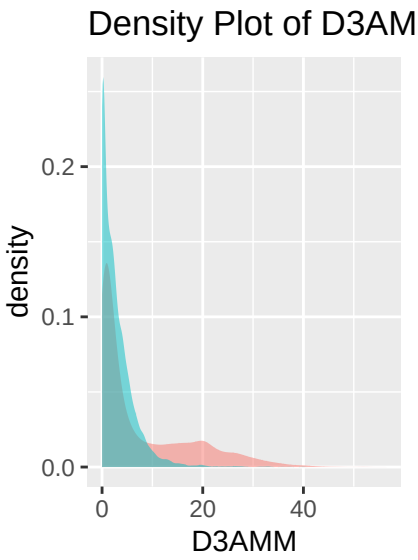


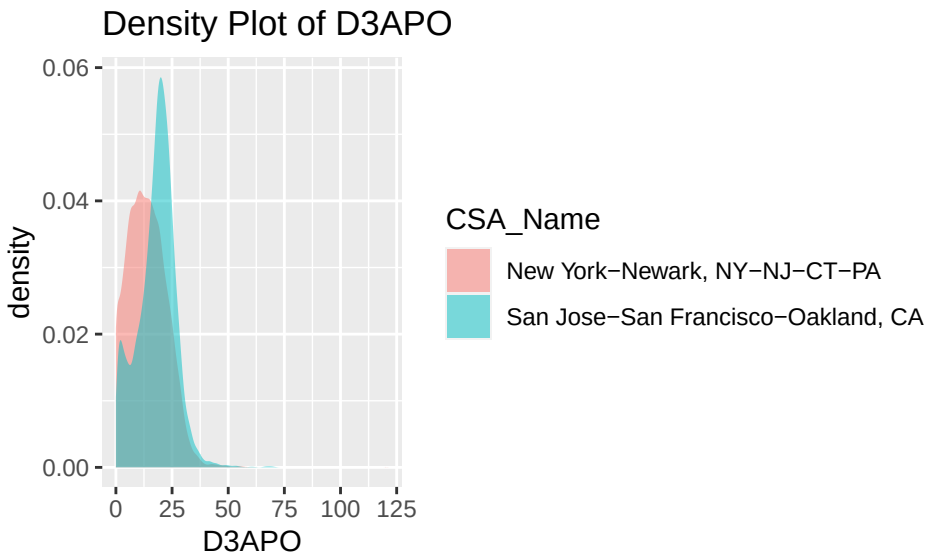


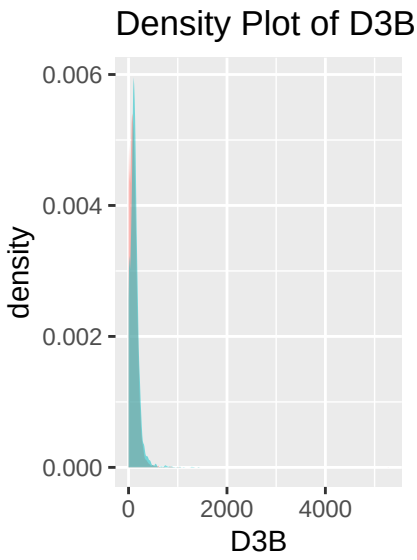

Density Plot of D2C_WREMLX
density
D2C_WREMLX
0.00
0.25
0.50
0.75
1.00
0
1
2
3
4
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

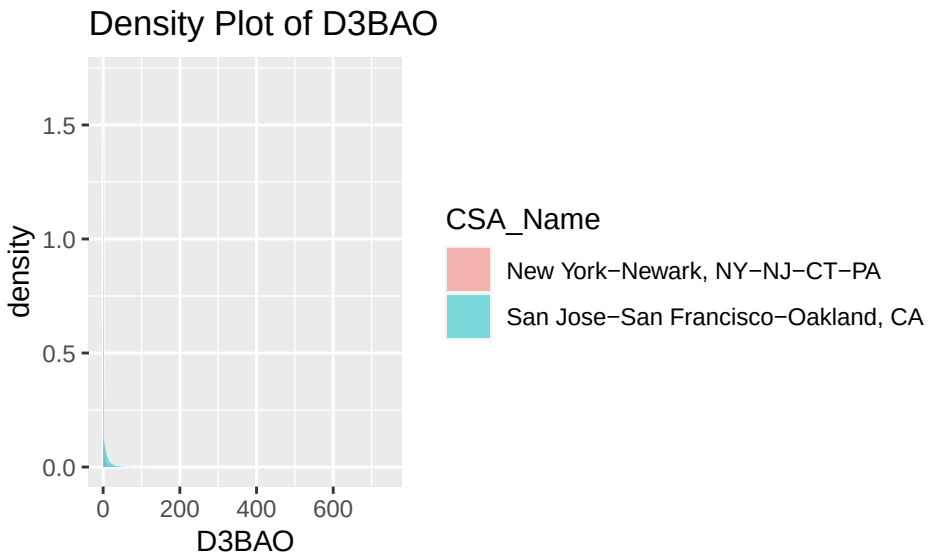

Density Plot of D3A
density
D3A
0
50
100
150
0.00
0.02
0.04

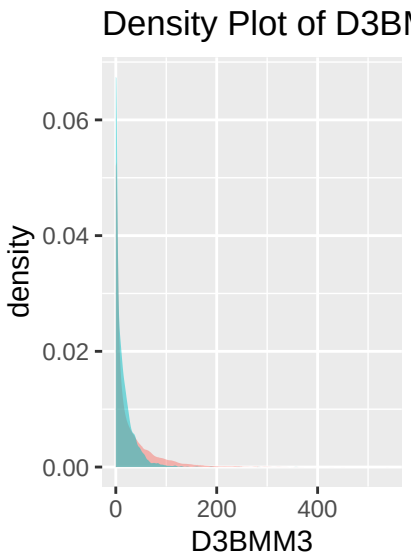

Density Plot of D3AAO
density
D3AAO
0
30
60
90
0.0
2.5
5.0
7.5
10.0
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

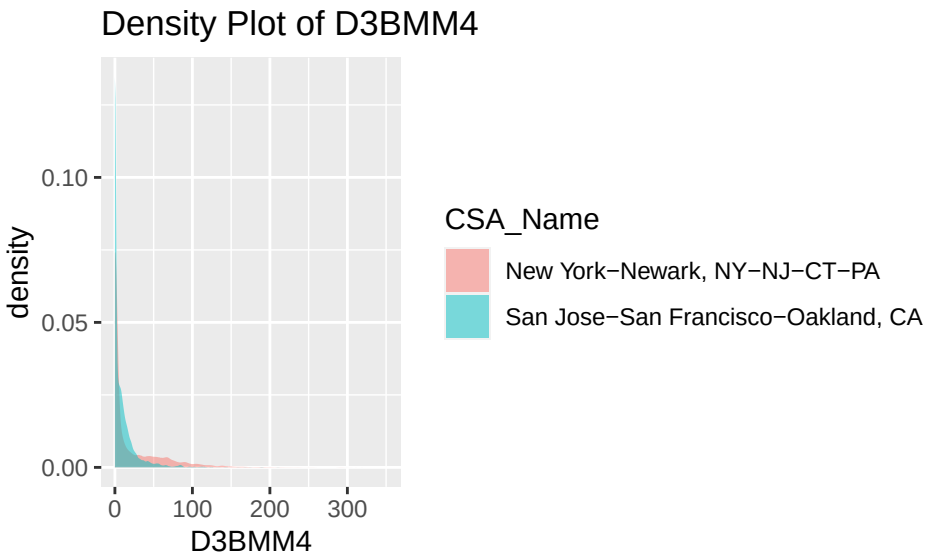

Density Plot of D3AMM
density
D3AMM
0
20
40
0.0
0.1
0.2

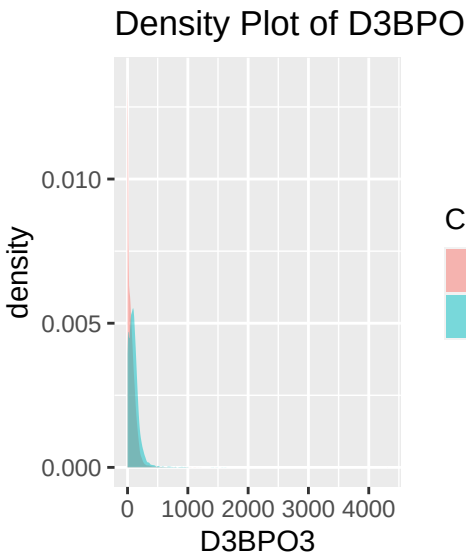

Density Plot of D3APO
density
0.06
0.04
0.02
0.00
0
25
50
75
100
125
D3APO
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

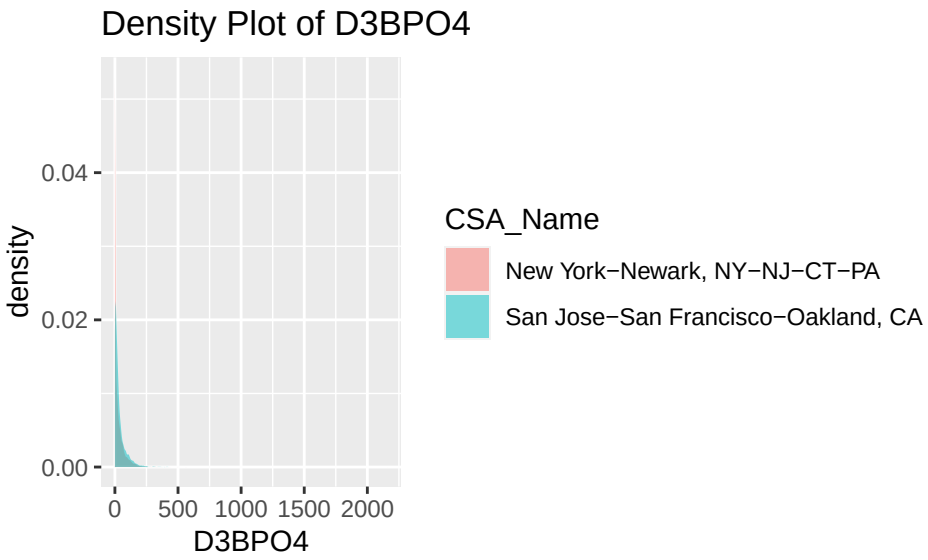

Density Plot of D3B
density
0.006
0.004
0.002
0.000
0
2000
4000
D3B

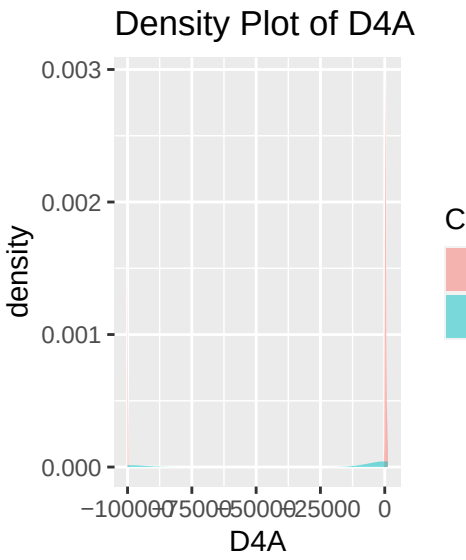

Density Plot of D3BAO
density
1.5
1.0
0.5
0.0
0
200
400
600
D3BAO
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

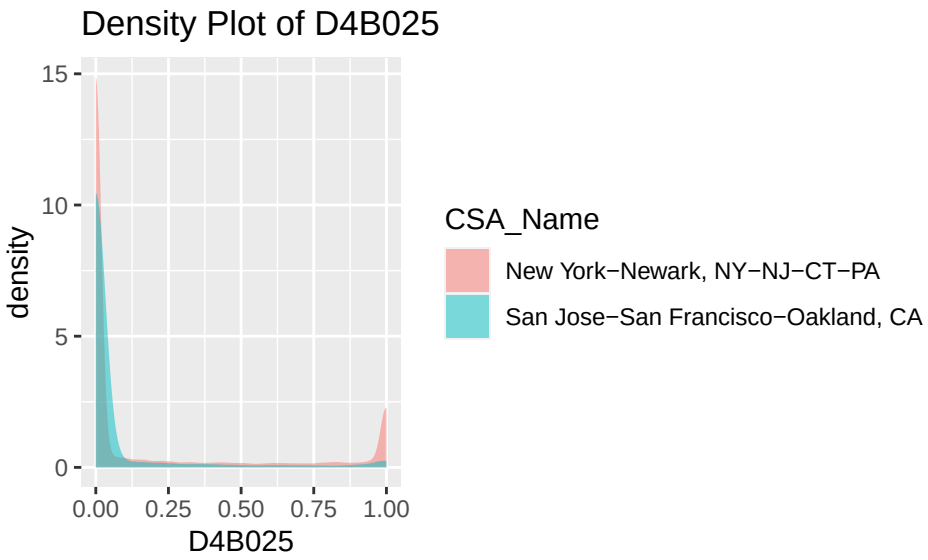

Density Plot of D3BMM3
density
0.06
0.04
0.02
0.00
0
200
400
D3BMM3

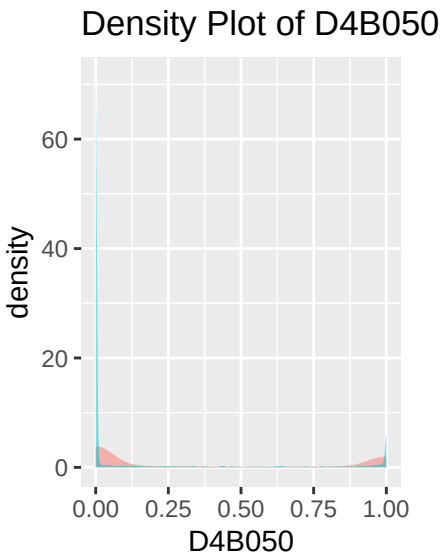

Density Plot of D3BMM4
density
0.10
0.05
0.00
0
100
200
300
D3BMM4
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

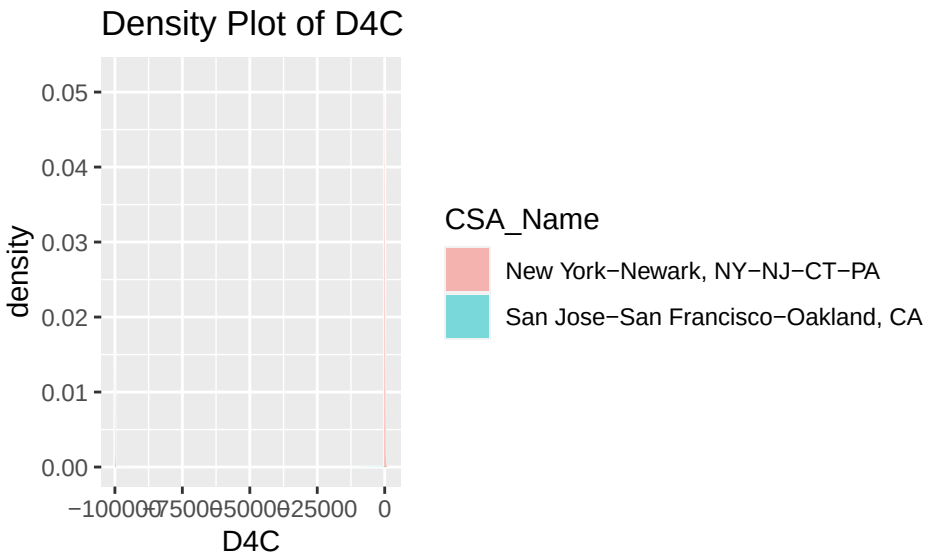

Density Plot of D3BPO
density
0.010
0.005
0.000
0
1000
2000
3000
4000
D3BPO3

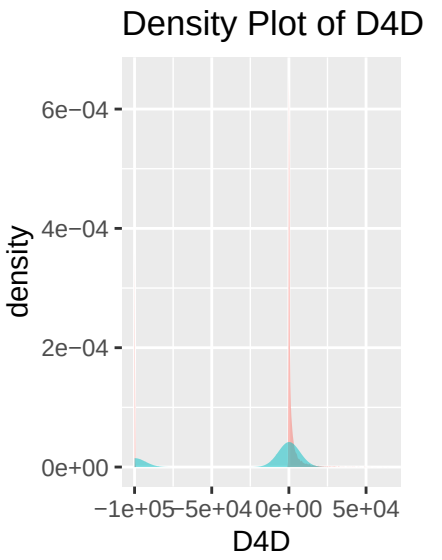

Density Plot of D3BPO4
density
0.04
0.02
0.00
0
500
1000
1500
2000
D3BPO4
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

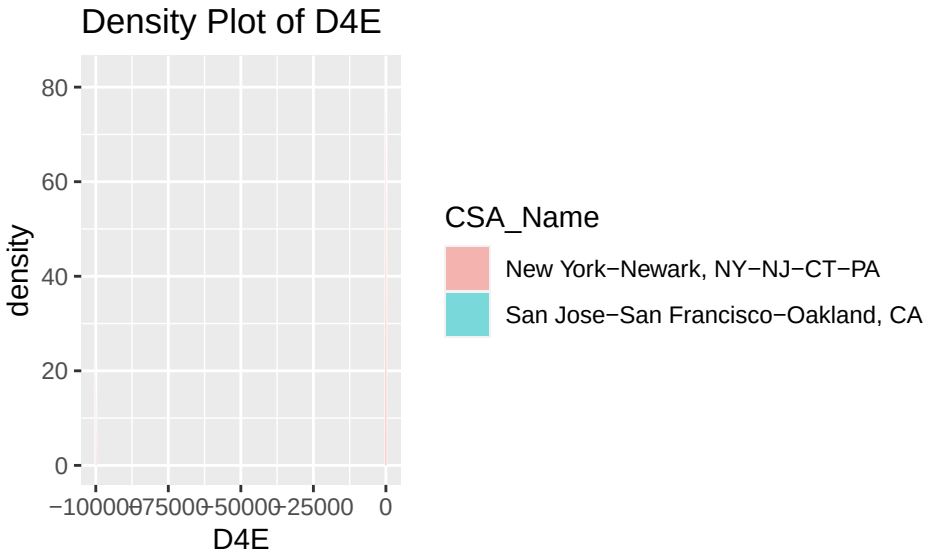

Density Plot of D4A
density
0.003
0.002
0.001
0.000
−100000
−75000
−50000
−25000
0
D4A

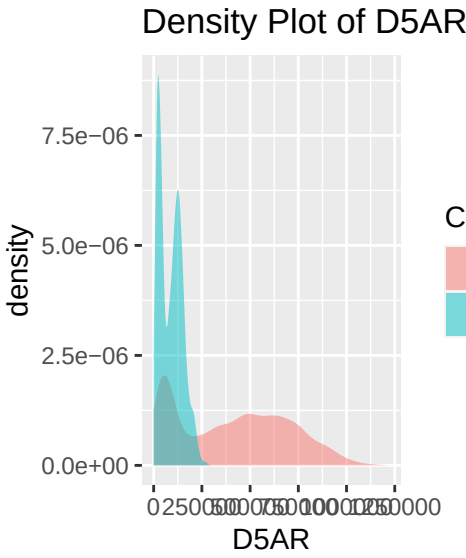

Density Plot of D4B025
density
D4B025
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

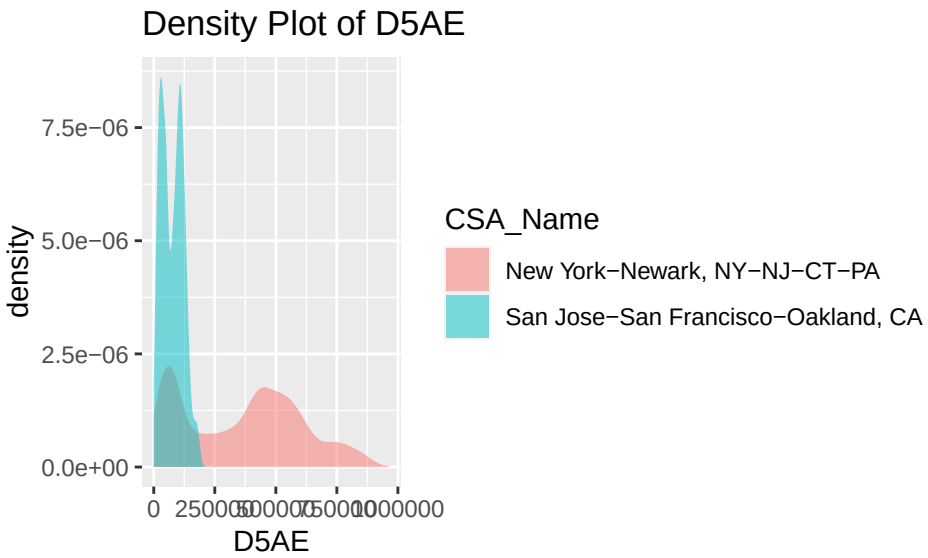

Density Plot of D4B050
density
D4B050

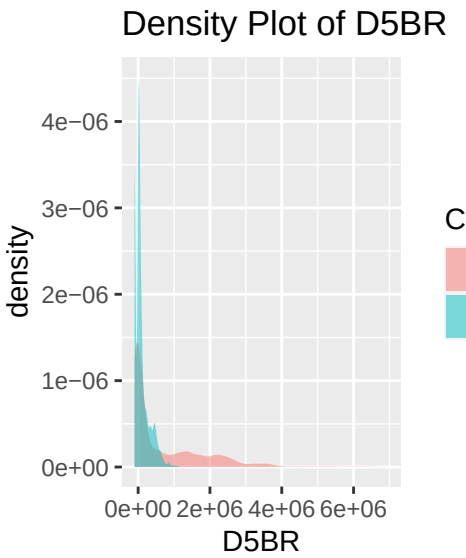

Density Plot of D4C
density
D4C
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

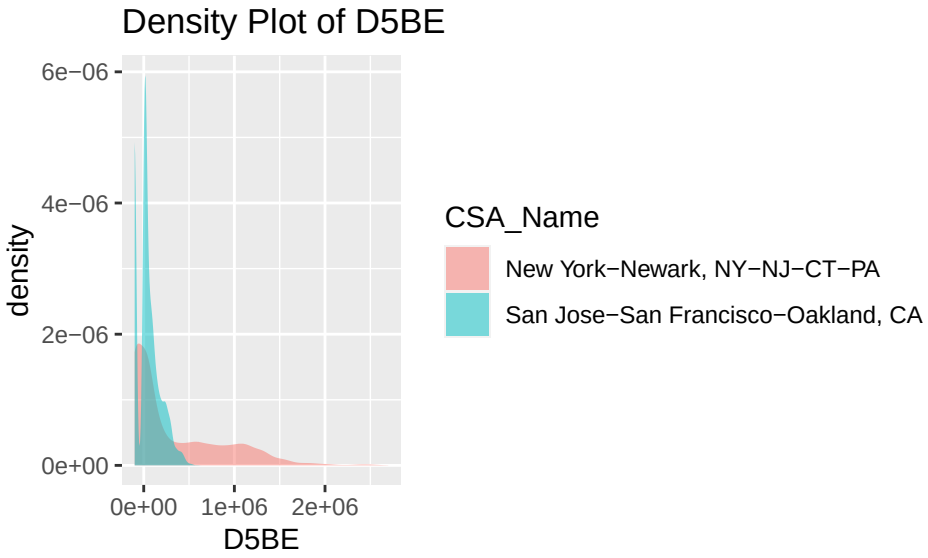

Density Plot of D4D
density
D4D

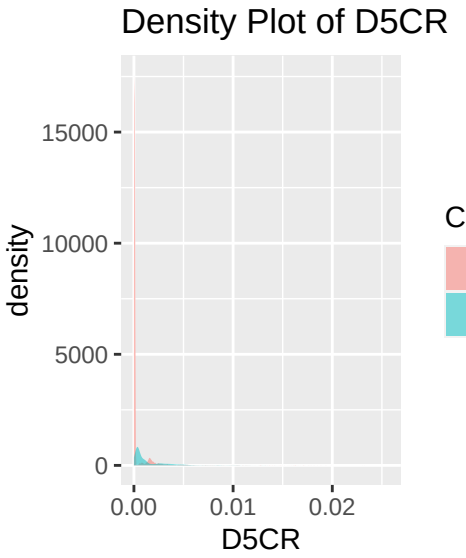

Density Plot of D4E
density
80
60
40
20
0
−100000
−75000
−50000
−25000
0
D4E
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

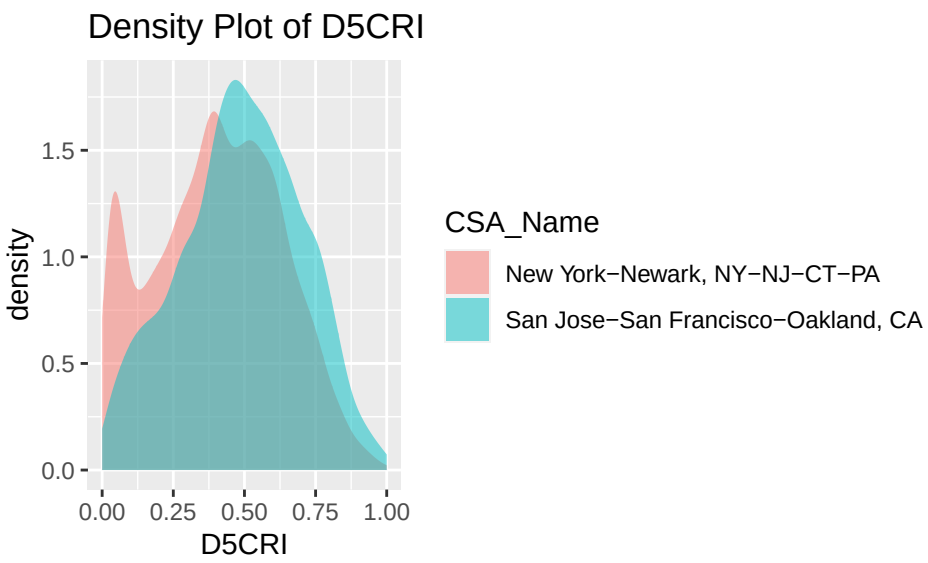

Density Plot of D5AR
density
7.5e−06
5.0e−06
2.5e−06
0.0e+00
D5AR

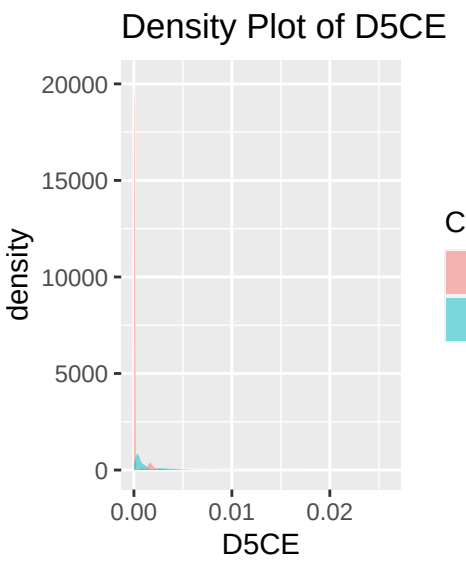

Density Plot of D5AE
density
7.5e−06
5.0e−06
2.5e−06
0.0e+00
0
D5AE
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

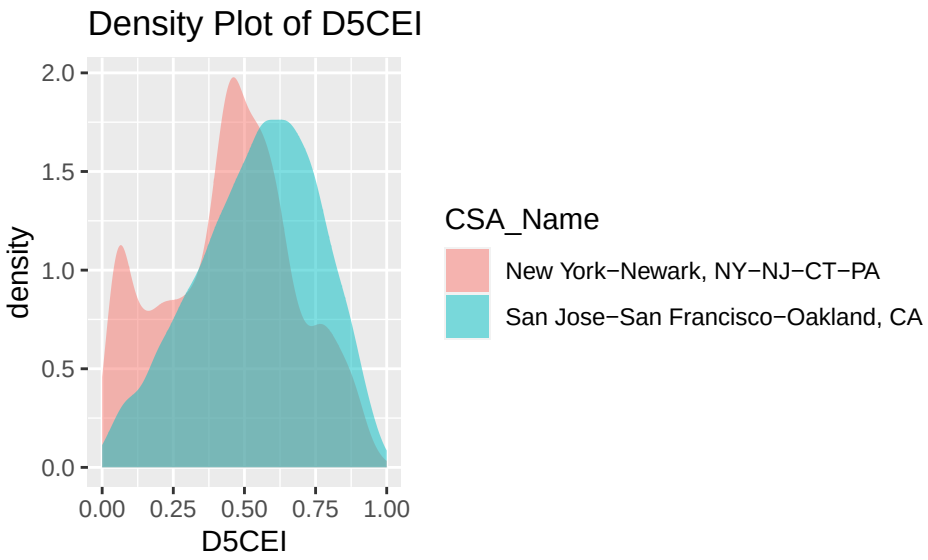

Density Plot of D5BR
density
4e−06
3e−06
2e−06
1e−06
0e+00
0e+00
2e+06
4e+06
6e+06
D5BR

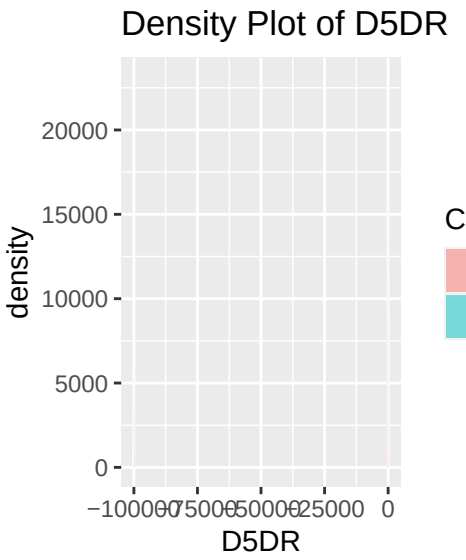

Density Plot of D5BE
6e−06
4e−06
2e−06
0e+00
density
0e+00
1e+06
2e+06
D5BE
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

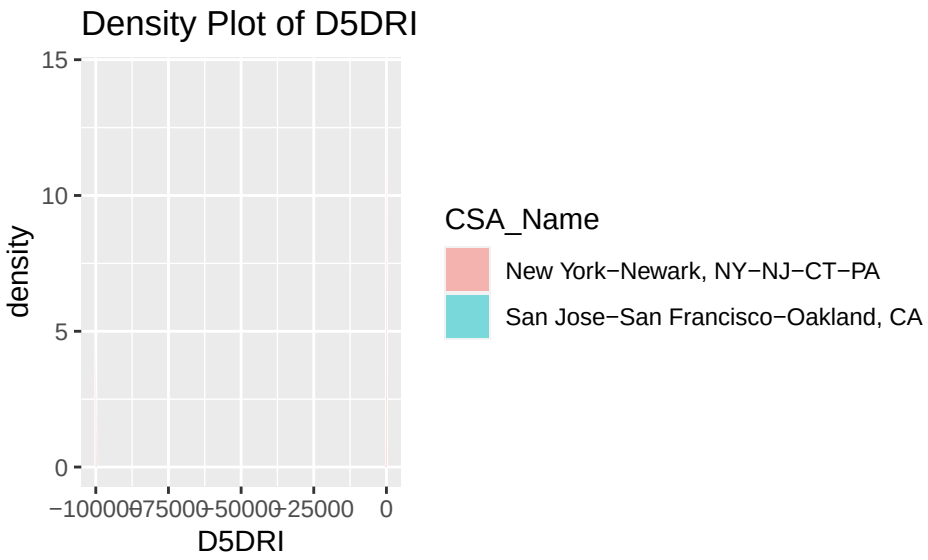

Density Plot of D5CR
15000
10000
5000
0
density
0.00
0.01
0.02
D5CR
C

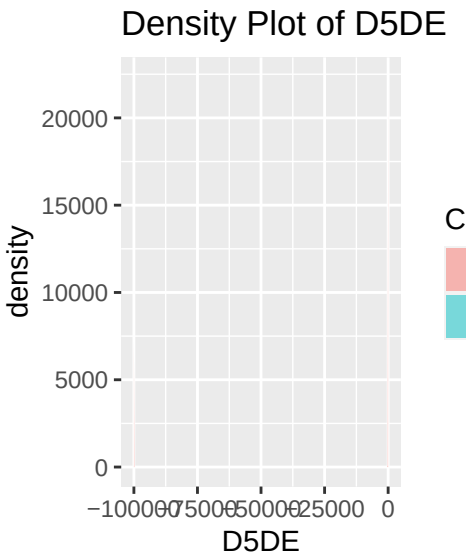

Density Plot of D5CRI
1.5
1.0
0.5
0.0
density
0.00
0.25
0.50
0.75
1.00
D5CRI
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

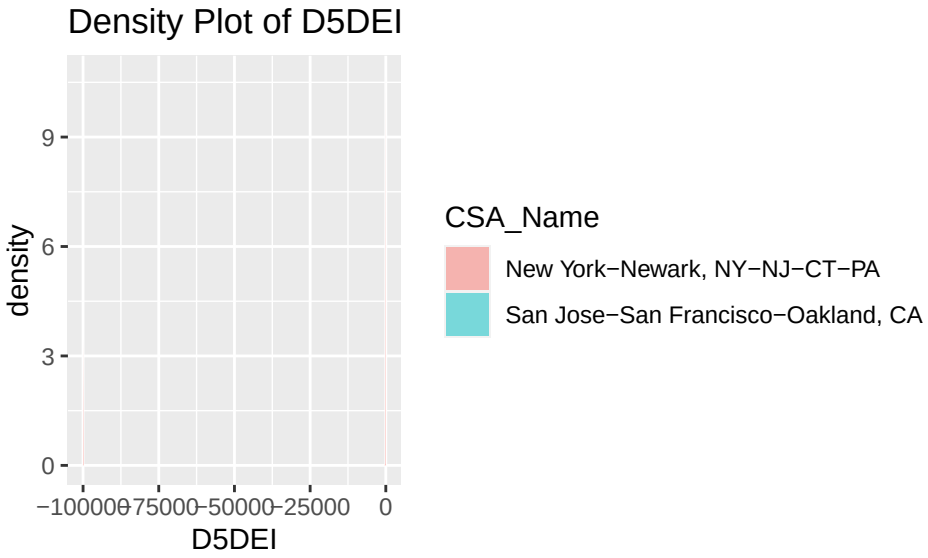

Density Plot of D5CE
20000
15000
10000
5000
0
density
0.00
0.01
0.02
D5CE
C

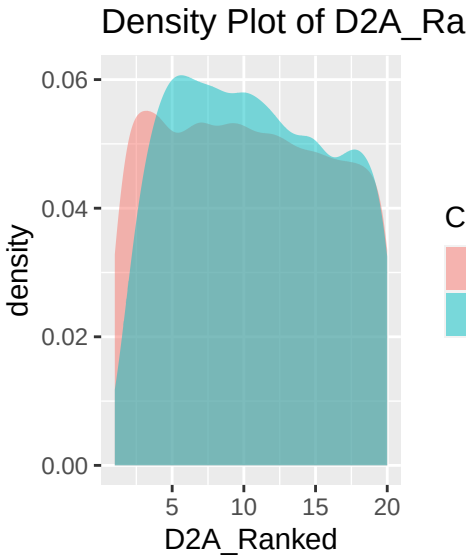

Density Plot of D5CEI
density
2.0
1.5
1.0
0.5
0.0
0.00
0.25
0.50
0.75
1.00
D5CEI
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

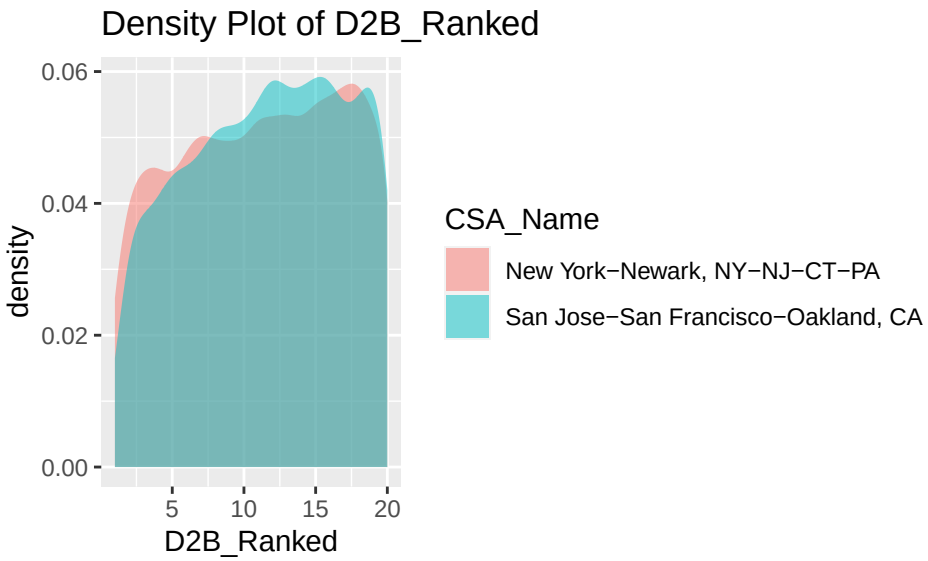

Density Plot of D5DR
density
20000
15000
10000
5000
0
−100000
−75000
−50000
−25000
0
D5DR
C

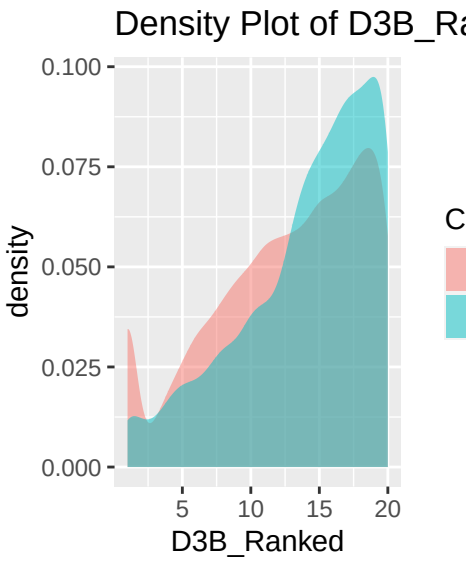

Density Plot of D5DRI
density
15
10
5
0
−100000
−75000
−50000
−25000
0
D5DRI
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

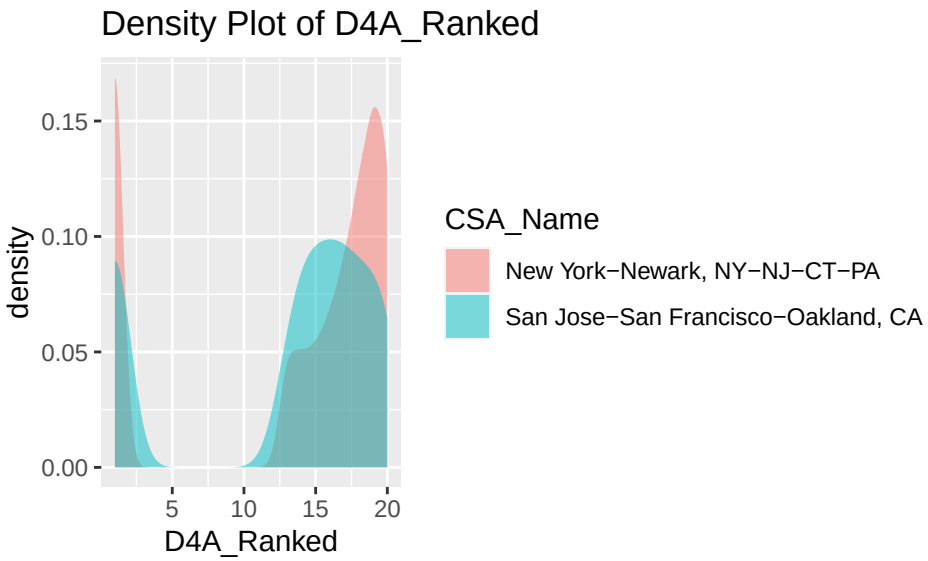

Density Plot of D5DE
density
20000
15000
10000
5000
0
−100000
−75000
−50000
−25000
0
D5DE
C

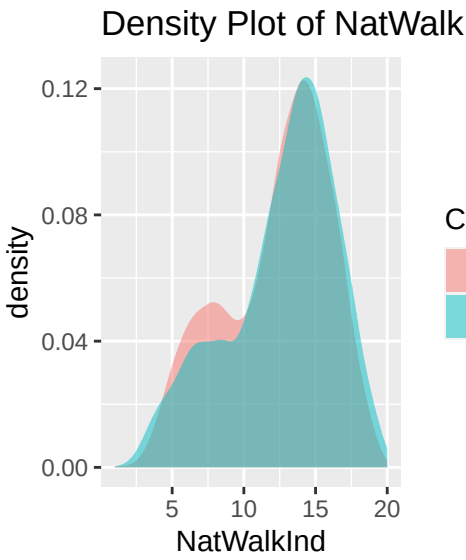

Density Plot of D5DEI
density
9
6
3
0
−100000
−75000
−50000
−25000
0
D5DEI
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA

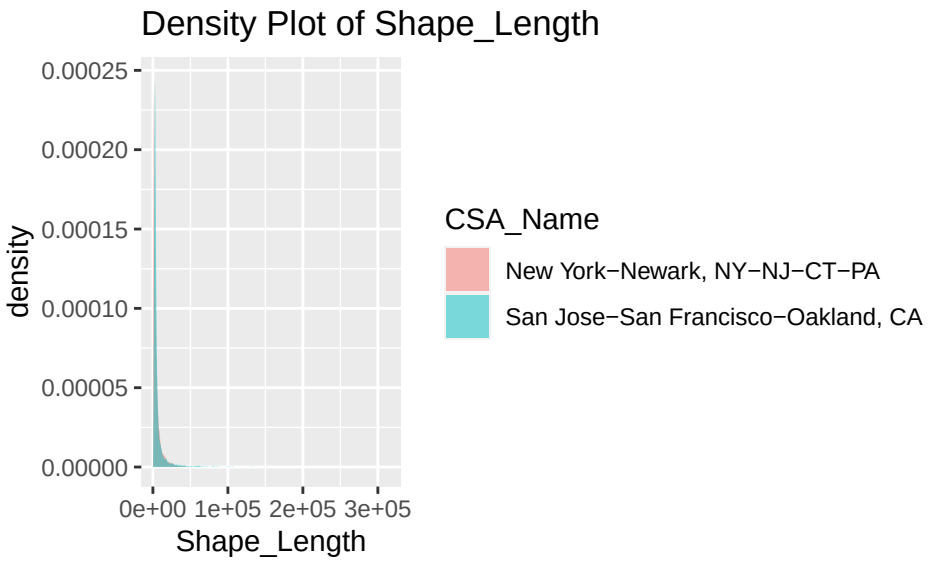

Density Plot of D2A_Ra
density
0.06
0.04
0.02
0.00
5
10
15
20
D2A_Ranked
C

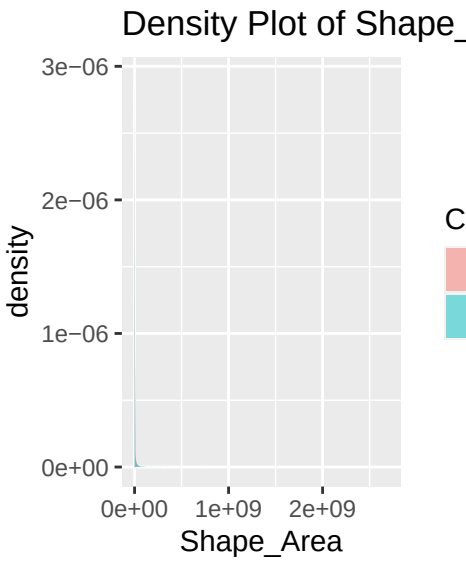

Density Plot of D2B_Ranked
density
0.06
0.04
0.02
0.00
5
10
15
20
D2B_Ranked
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA


Density Plot of D3B_Ra
density
0.100
0.075
0.050
0.025
0.000
5
10
15
20
D3B_Ranked
C

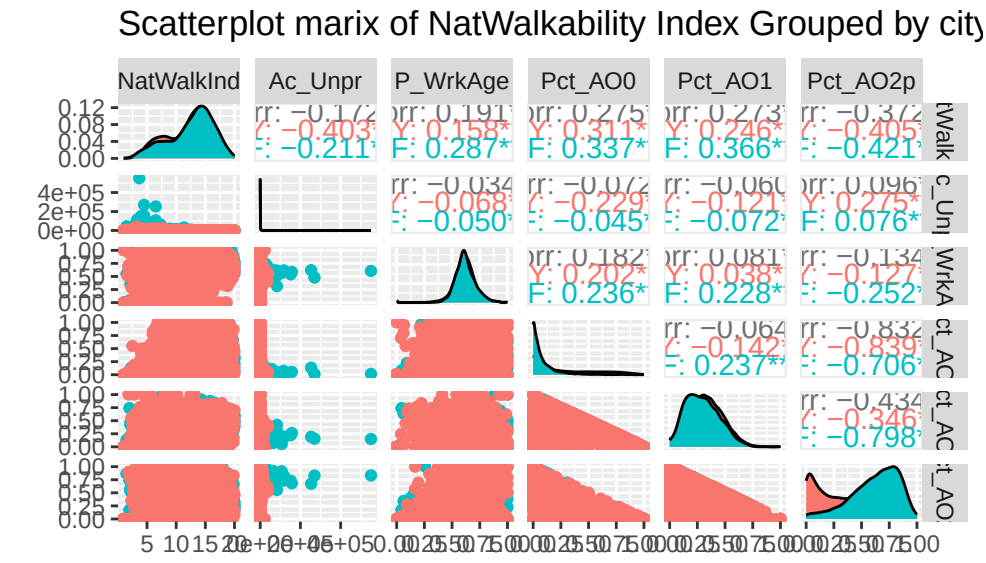

Density Plot of D4A_Ranked
density
0.15
0.10
0.05
0.00
5
10
15
20
D4A_Ranked
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA


Density Plot of NatWalk
density
0.12
0.08
0.04
0.00
5
10
15
20
NatWalkInd
C

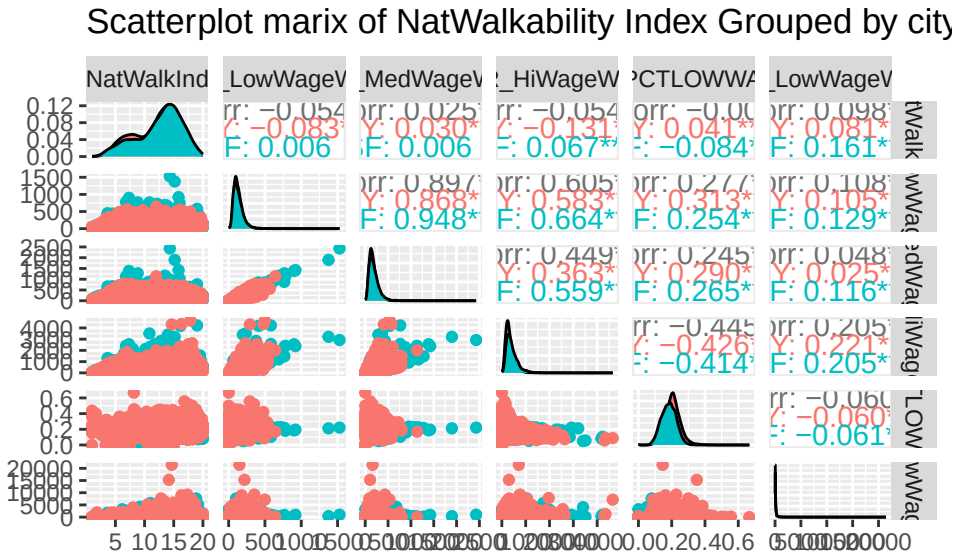

Density Plot of Shape_Length
density
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000
0e+00
1e+05
2e+05
3e+05
Shape_Length
CSA_Name
New York−Newark, NY−NJ−CT−PA
San Jose−San Francisco−Oakland, CA


Density Plot of Shape_
density
3e−06
2e−06
1e−06
0e+00
0e+00
1e+09
2e+09
Shape_Area
C

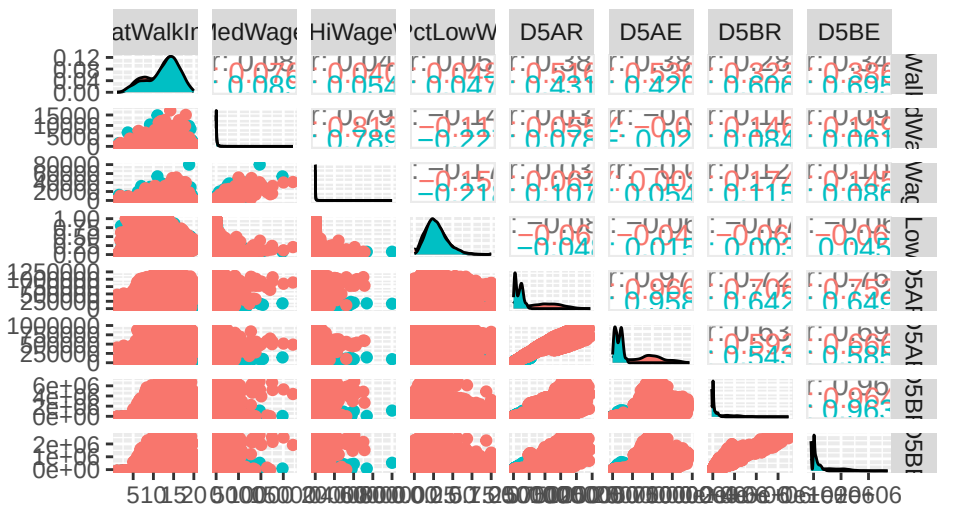

```
plot_1 <- suppressWarnings(ggpairs(NY_SF_focus[, c(1, 3:7)],
        aes(color = NY_SF_focus$CSA_Name_ABBREV),
        title = "Scatterplot marix of NatWalkability Index Grouped by city",
        subtitle = "Using the first 4 potential predictors"))
plot_1
```




```
ggpairs(NY_SF_focus[, c(1, 8:12)],
        aes(color = NY_SF_focus$CSA_Name_ABBREV),
        title = "Scatterplot marix of NatWalkability Index Grouped by city",
        subtitle = "Using the next 5 potential predictors")
```




```
ggpairs(NY_SF_focus[, c(1, 13:19)],
        aes(color = NY_SF_focus$CSA_Name_ABBREV),
        title = "Scatterplot marix of NatWalkability Index Grouped by city",
        subtitle = "Using the following 5 potential predictors")
```




```
ggpairs(NY_SF_focus[, c(1, 20:24)],
        aes(color = NY_SF_focus$CSA_Name_ABBREV),
        title = "Scatterplot marix of NatWalkability Index Grouped by city",
        subtitle = "Using the last 4 potential predictors")
```

Scatterplot marix of NatWalkability Index Grouped by city