Analysis of Drugnet network

CSS 692 – Final Presentation

Ajay Kulkarni

G01024139

Objectives

- Identification of dominant people to reduce the spread of drugs and diseases
- Exploration of local network structures which forms the global pattern
- Comparison of observed network and simulated network

Introduction

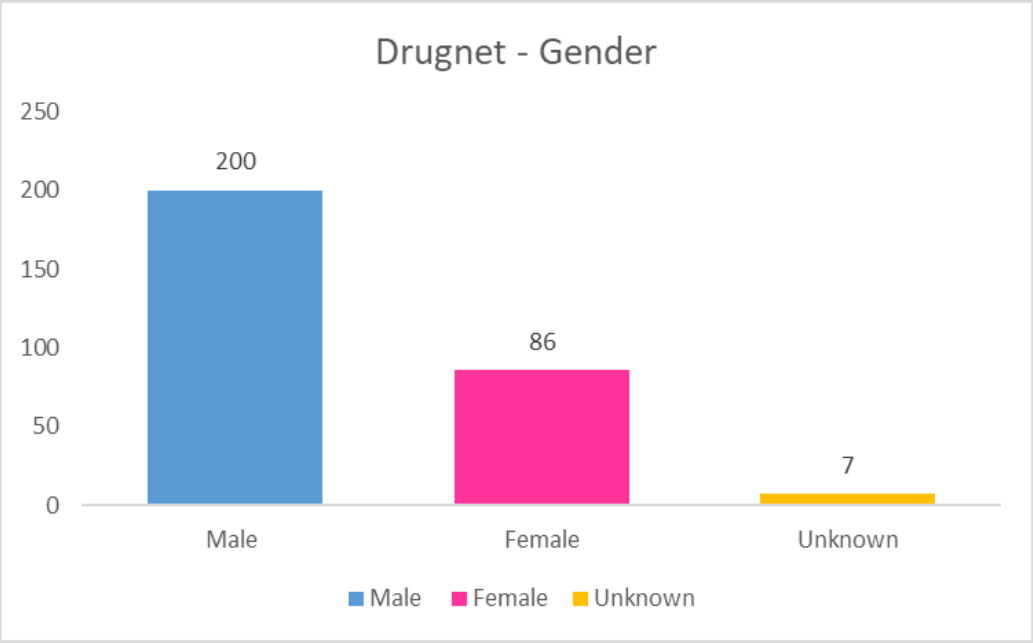
- Network represents the interaction of 293 people from Hartford, CT in the form of an adjacency matrix

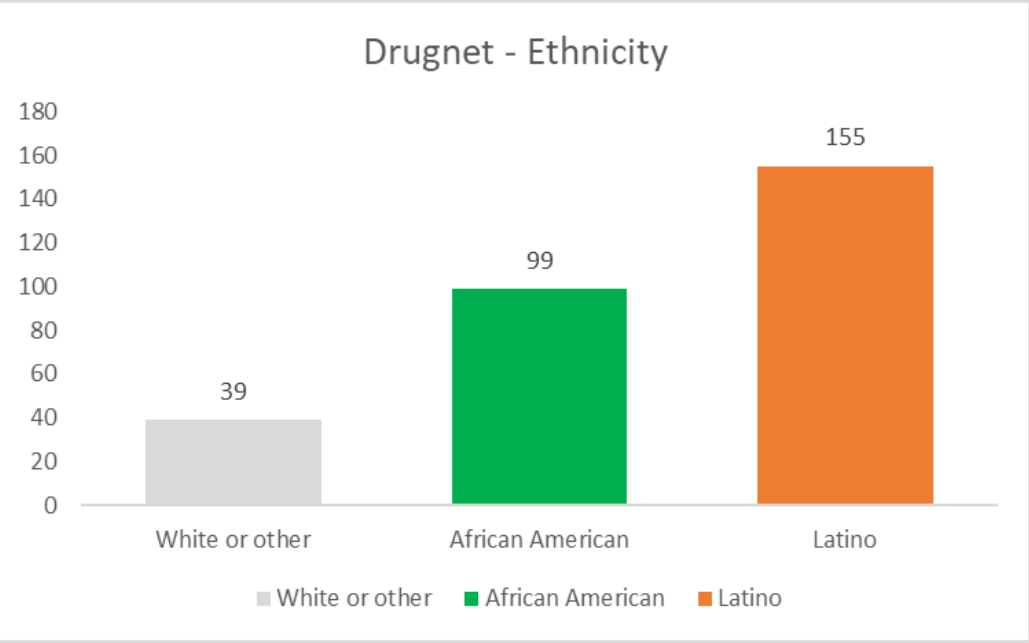
- The survey sample constructed through two methods
  - The majority (55%) was recruited through street outreach
  - The rest of the cohort was referred by the survey participants

- Eligibility criteria
  - At least 18 years of age
  - Reported active use of heroin, cocaine/crack or other injected illicit drug

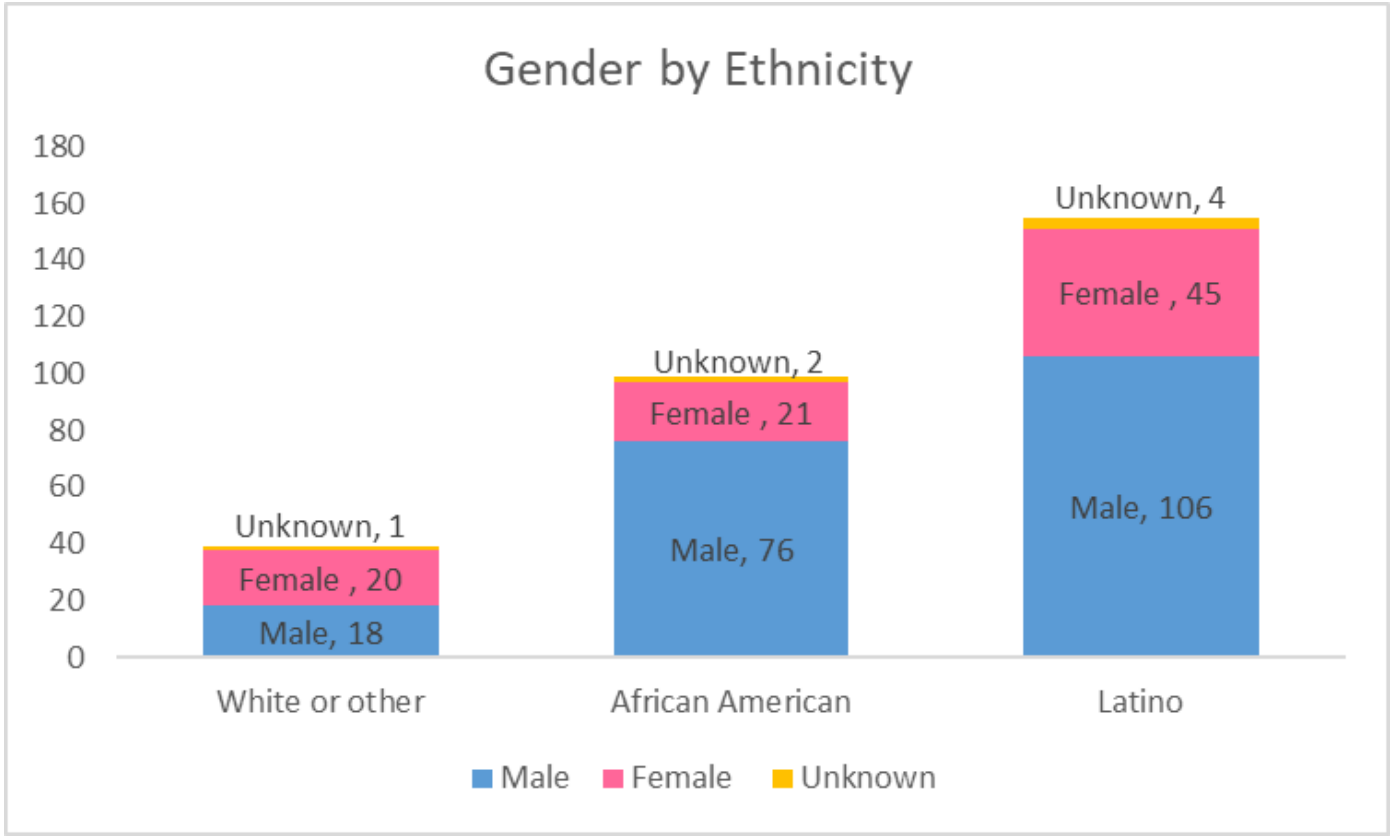
Network attributes

- Network has information about two attributes – Gender and Ethnicity

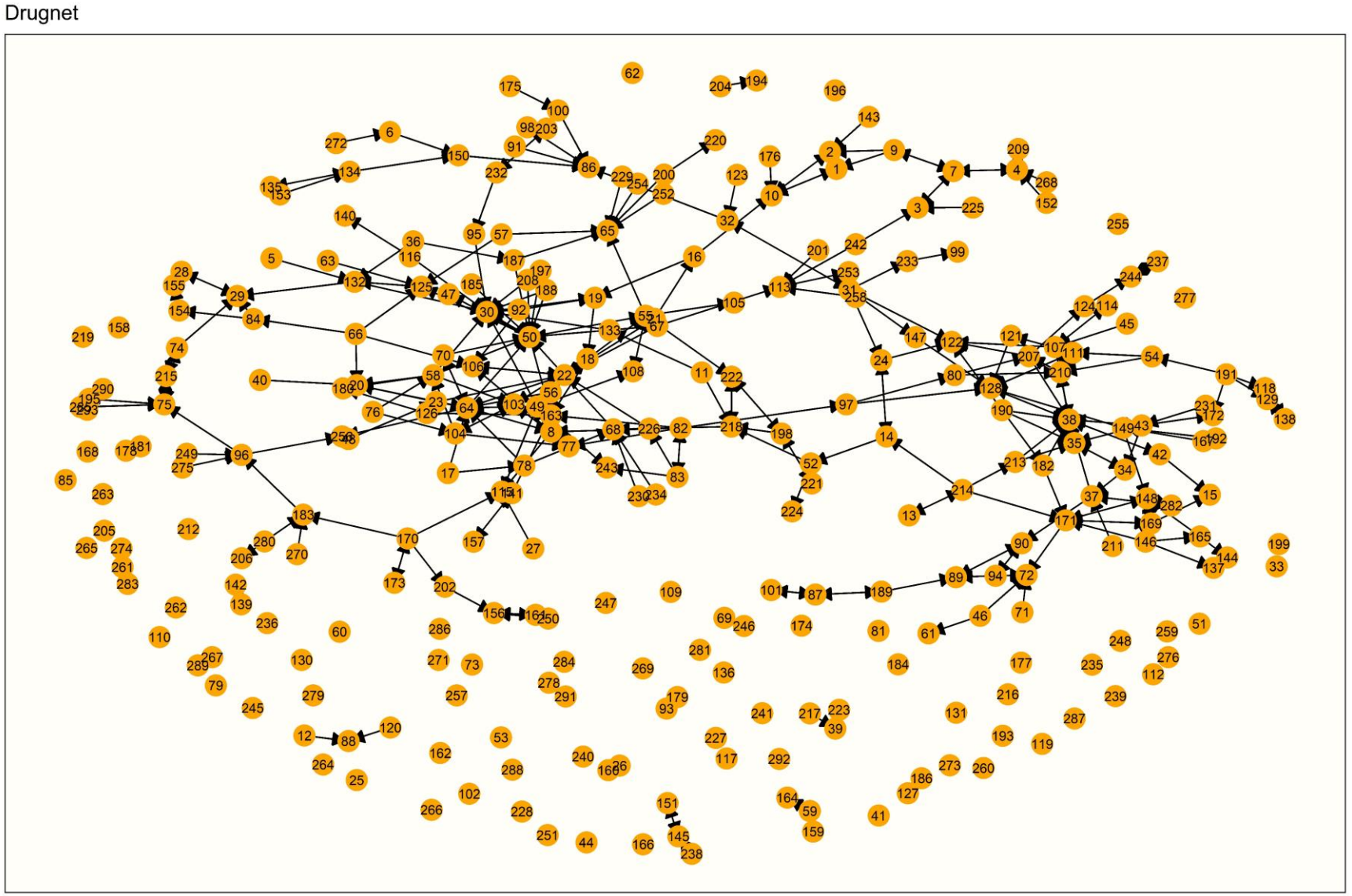




Network attributes

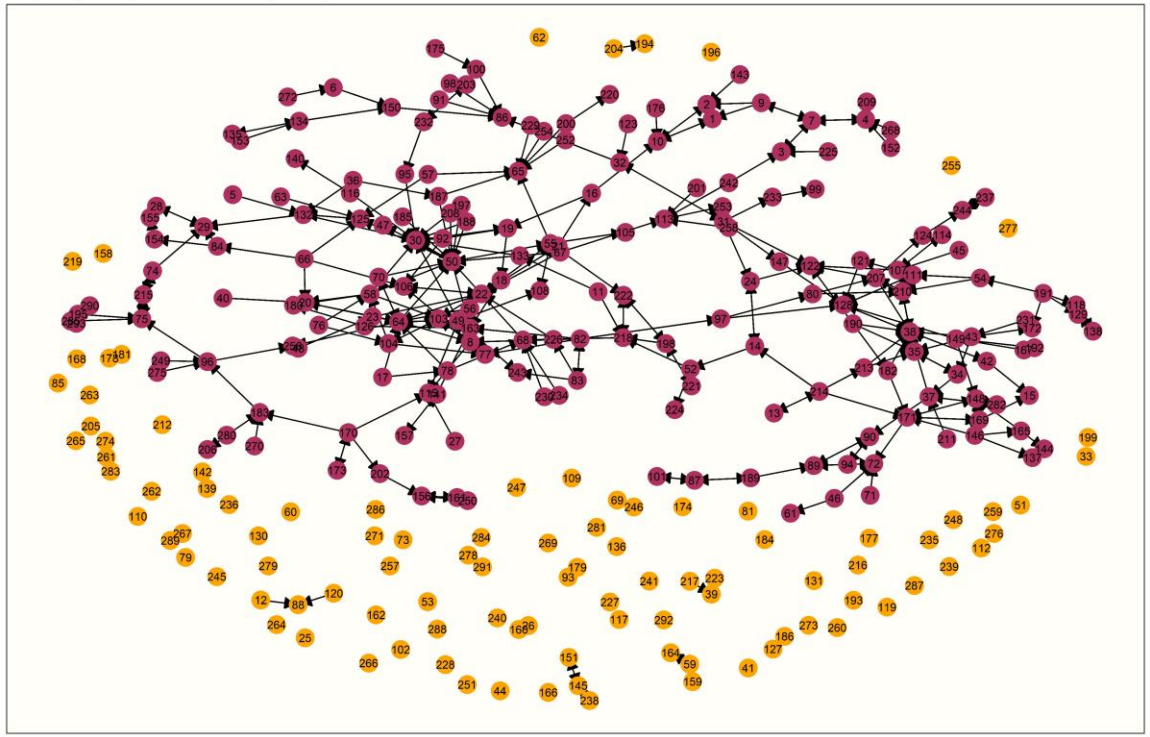


Network structure

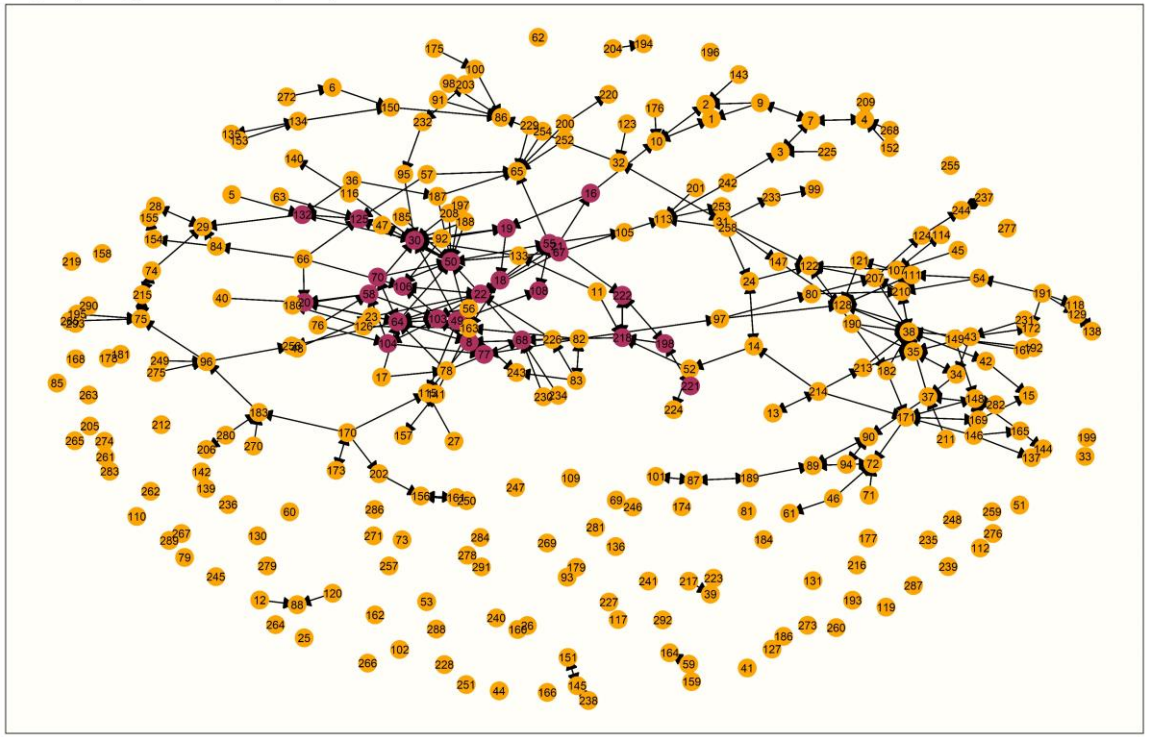


Network structure

Drugnet (Weakly Connected Component)



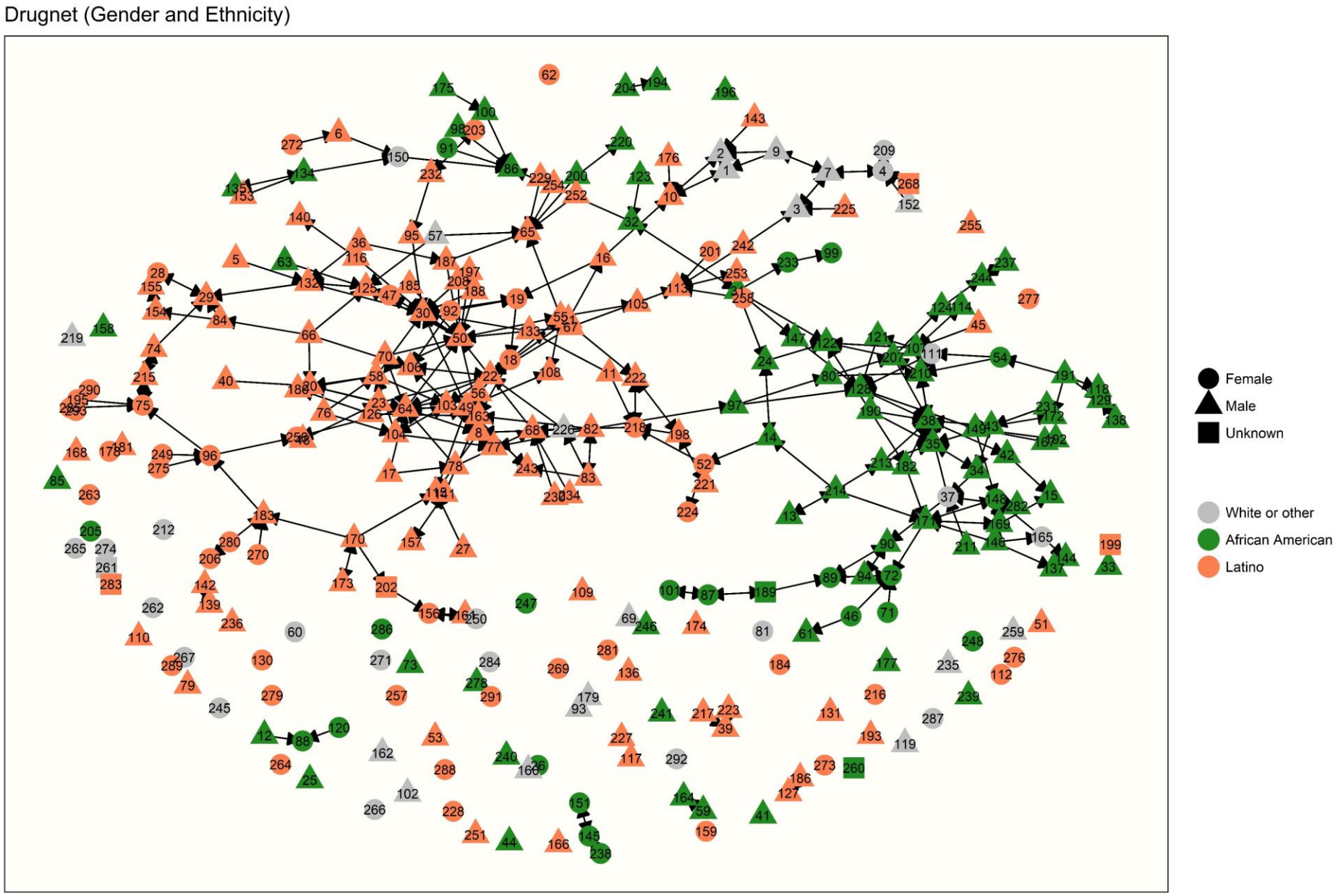
Drugnet (Strongly Connected Component)



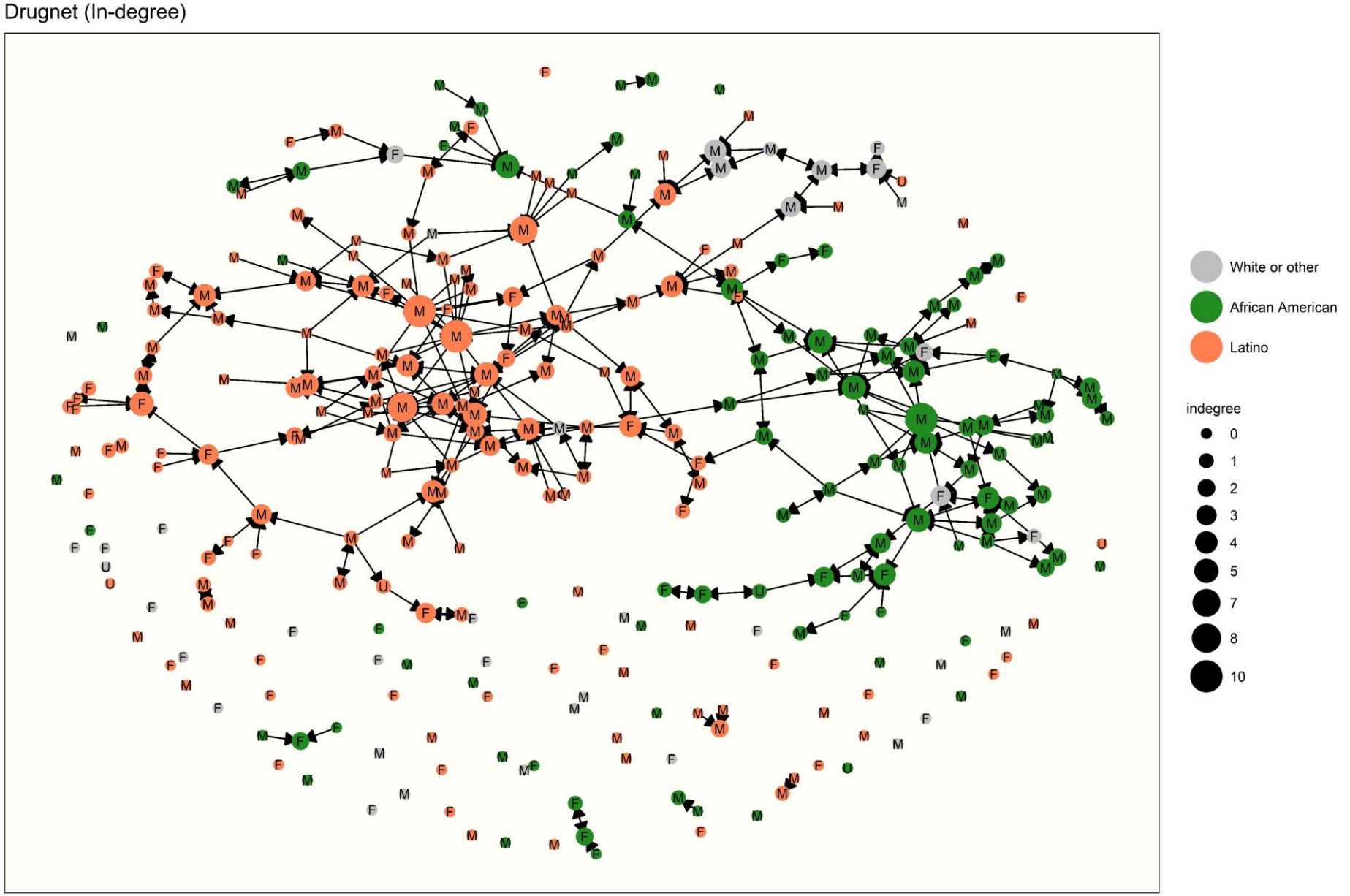
- Weakly connected component
  - Number of nodes – 193
  - Gender -> Male – 143 and Female – 43
  - Ethnicity -> White or other - 15
    African American - 69
    Latino - 109

- Strongly connected component
  - Number of nodes – 27
  - Gender -> Male – 24 and Female – 3
  - Ethnicity -> Latino - 27

Network structure - Gender and Ethnicity



In-degree centrality



In-degree centrality

In-degree	Node id	Gender	Ethnicity
10	30	Male	Latino
	38	Male	African American
	50	Male	Latino
8	64	Male	Latino
7	65	Male	Latino
5	22	Male	Latino
	75	Female	Latino
	87	Male	African American
	124	Male	African American
	130	Male	African American
	165	Male	Latino
	173	Male	African American

ERGMs

- ERGMs help us to understand about local processes and their interaction to form global network patterns

- In ERGM we include a parameter for each configuration and we estimate the value of the parameter that best matches the observed network

- These coefficients are estimated using MCMC (Markov Chain Monte Carlo) technique

ERGMs

- Three different models were built
  - Erdos-Renyi model (p_0 model) -> Edges
  - Holland and Leinhardt's p1 model -> Edges, Mutuality, Sender's effect & Receiver's effect
  - Created a model by selecting attributes based on Morris et al. (2008)

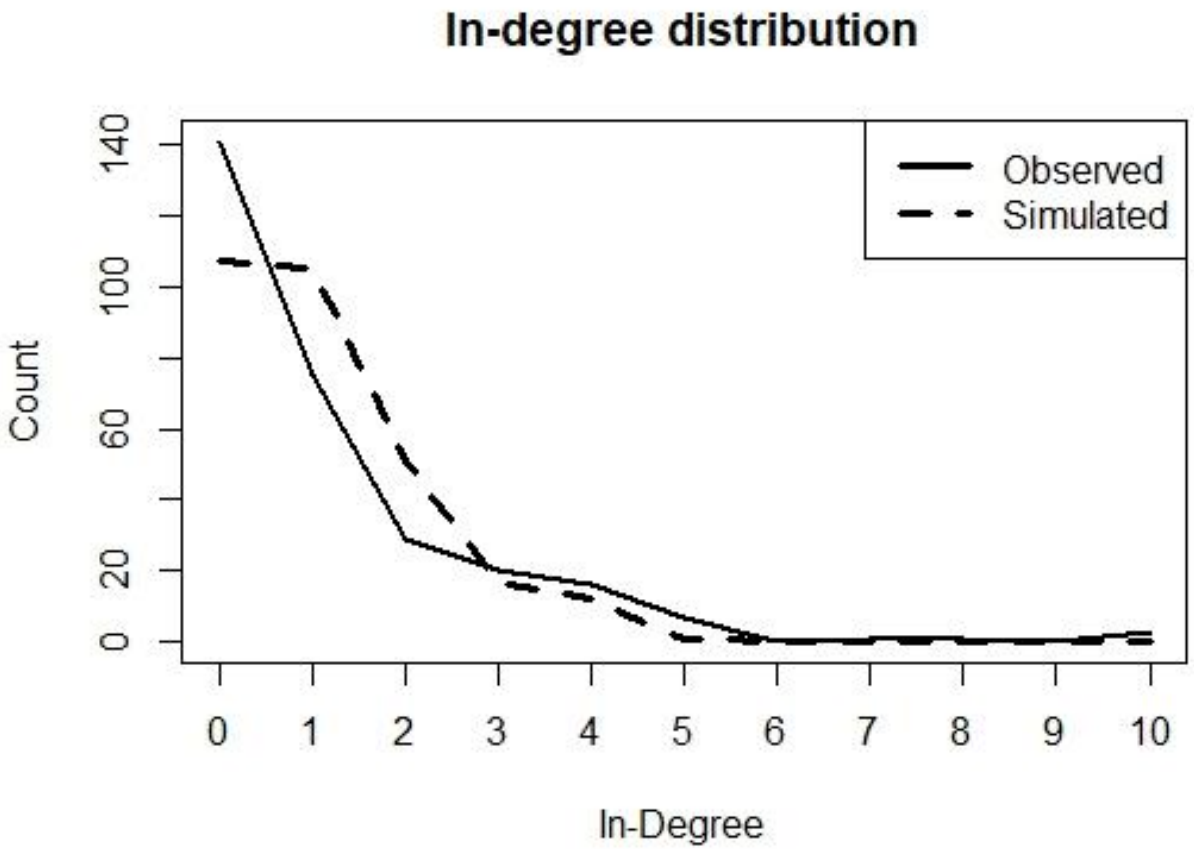
Erdos-Renyi model (p_0 model)

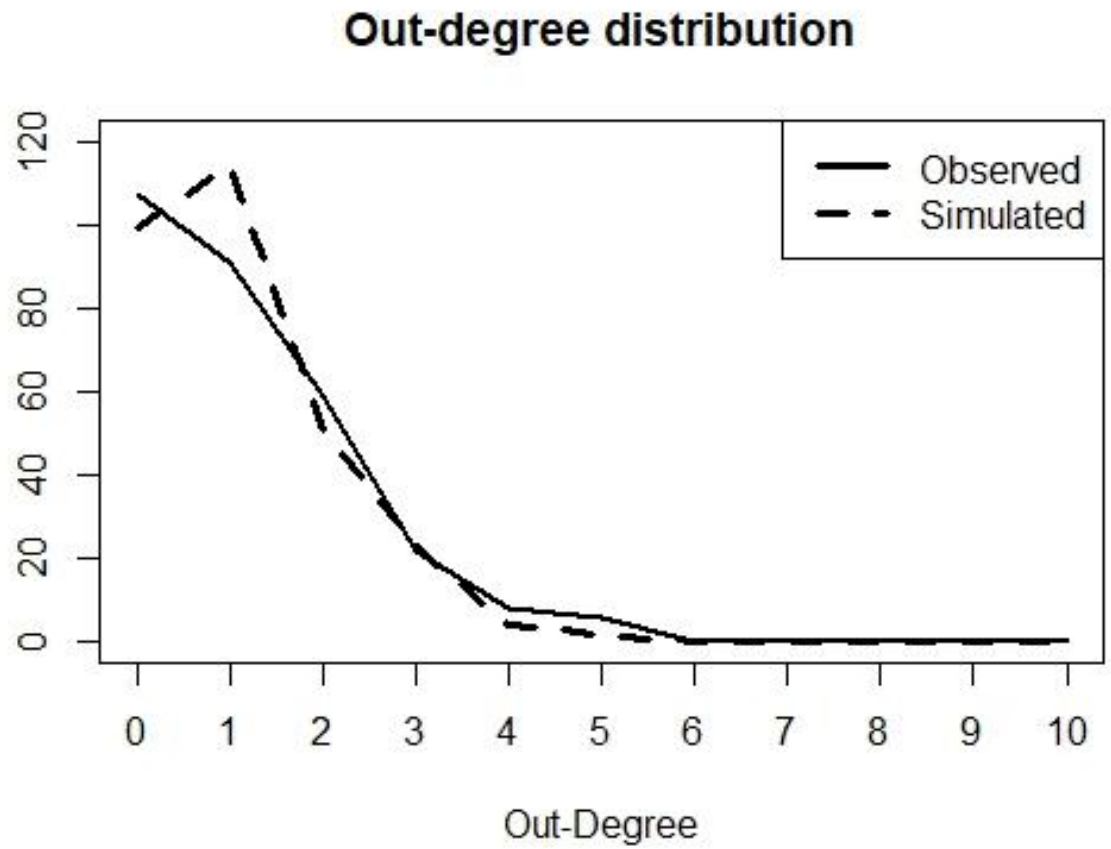
Console Terminal ×

C:/Users/ajask/Downloads/Social Network Analysis/Drugnet CSV/

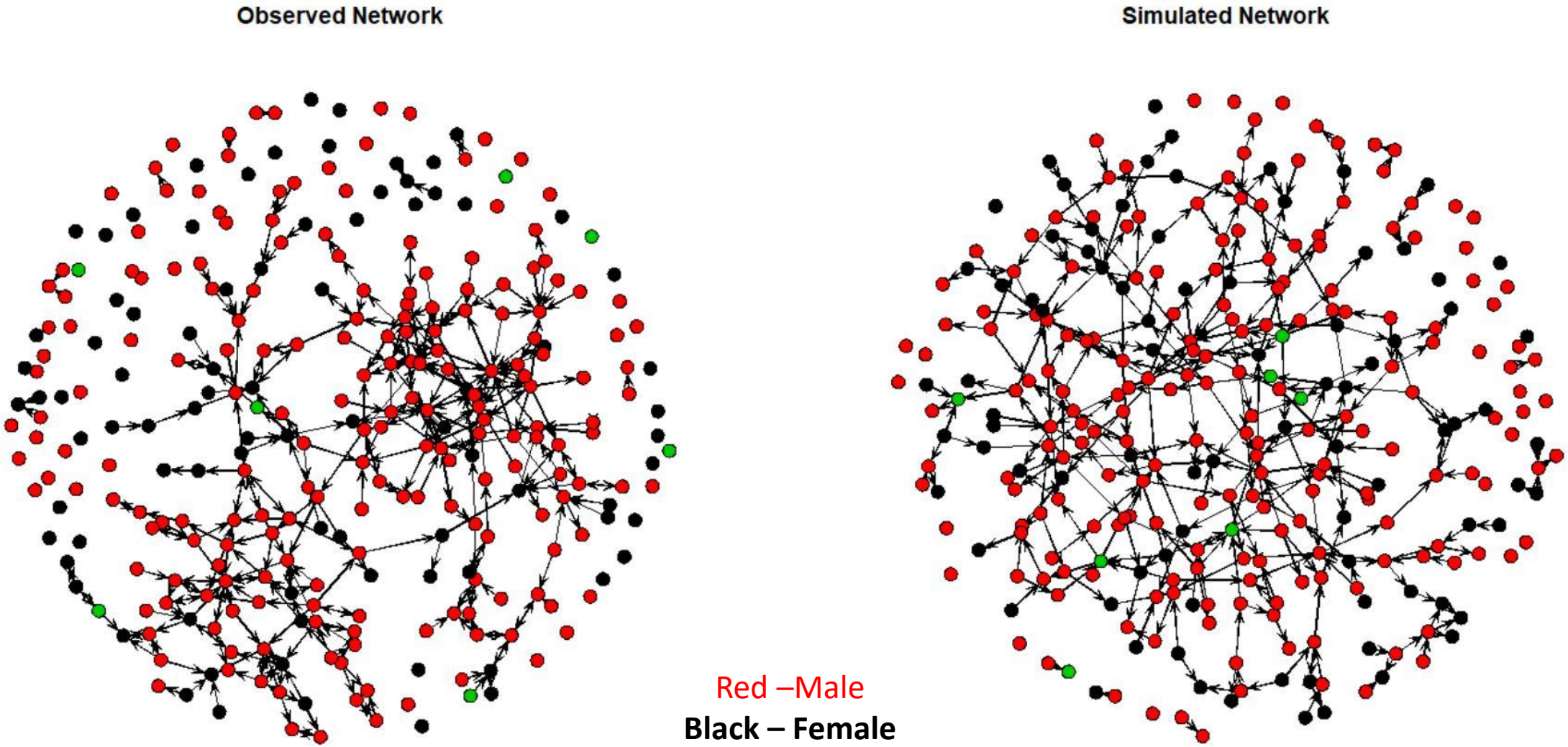
```
> m1 = ergm(net ~ edges)
Evaluating log-likelihood at the estimate.
> summary(m1)

==========================
Summary of model fit
==========================

Formula:   net ~ edges

Iterations:  8 out of 20

Monte Carlo MLE Results:
      Estimate Std. Error MCMC % p-value
edges -5.53290    0.05458      0  <1e-04 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

     Null Deviance: 118606  on 85556  degrees of freedom
 Residual Deviance:   4405  on 85555  degrees of freedom

AIC: 4407    BIC: 4416    (Smaller is better.)
>
```

Erdos-Renyi model (p_0 model)





Erdos-Renyi model (p_0 model)



Red –Male

Black – Female

Green - Unknown

Model based on selected attributes

- Model was built by selecting statistically relevant attributes from the list provided by Morris et al. (2008)

- Selected attributes

Attribute	Explanation
Edges	Edges in the network
Mutual	Mutuality
Asymmetric	Asymmetric dyads
Nodefactor	Main effect of a factor attribute
Nodematch	Interaction term - homophily
Simmelianties	Ties in Simmelian triads
Isolates	Isolates in the network

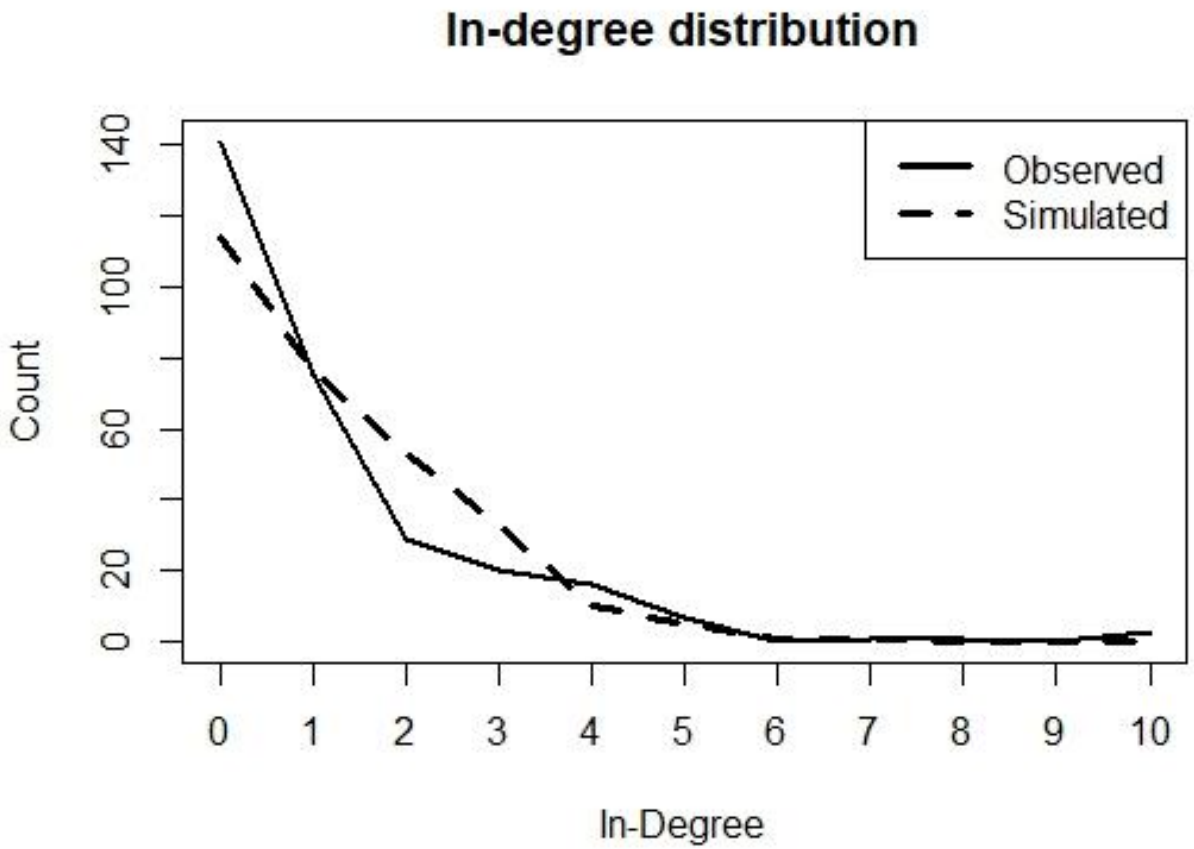
Model based on selected attributes

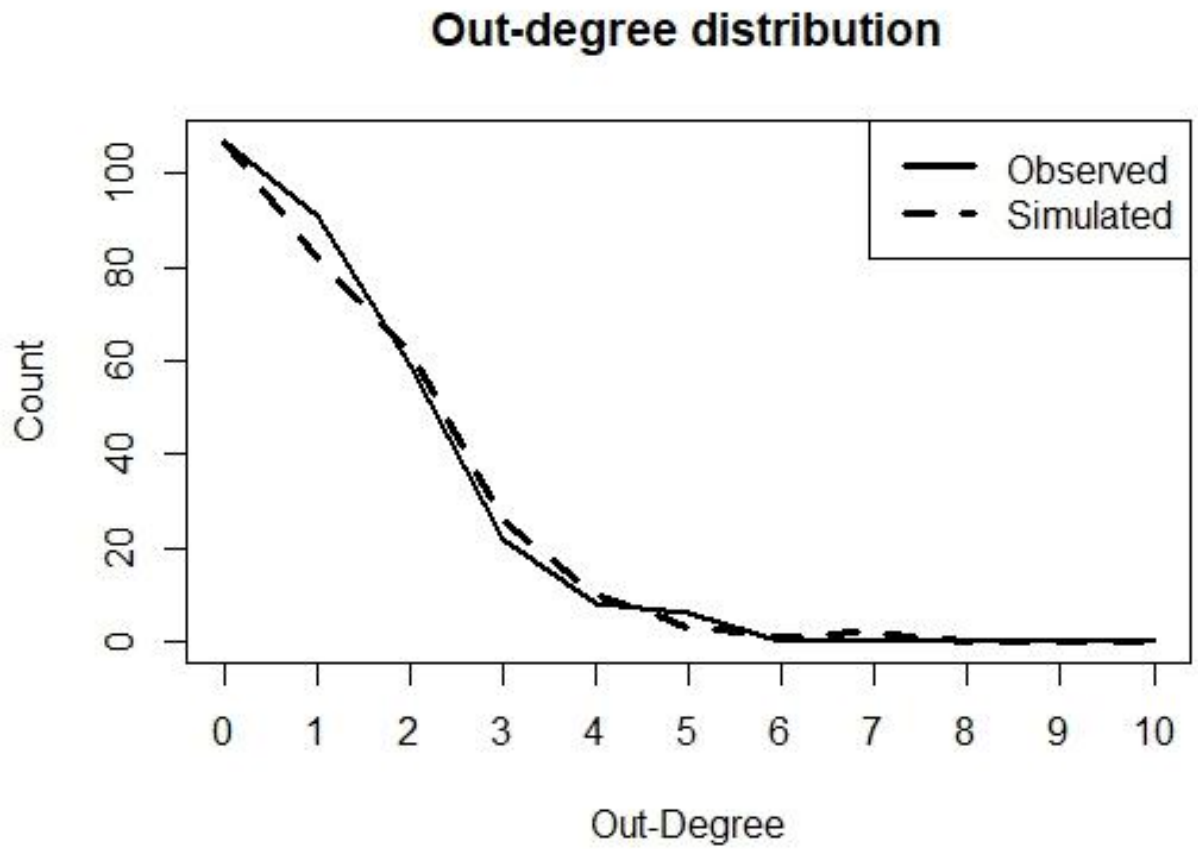
Console Terminal

C:/Users/ajask/Downloads/Social Network Analysis/Drugnet CSV/

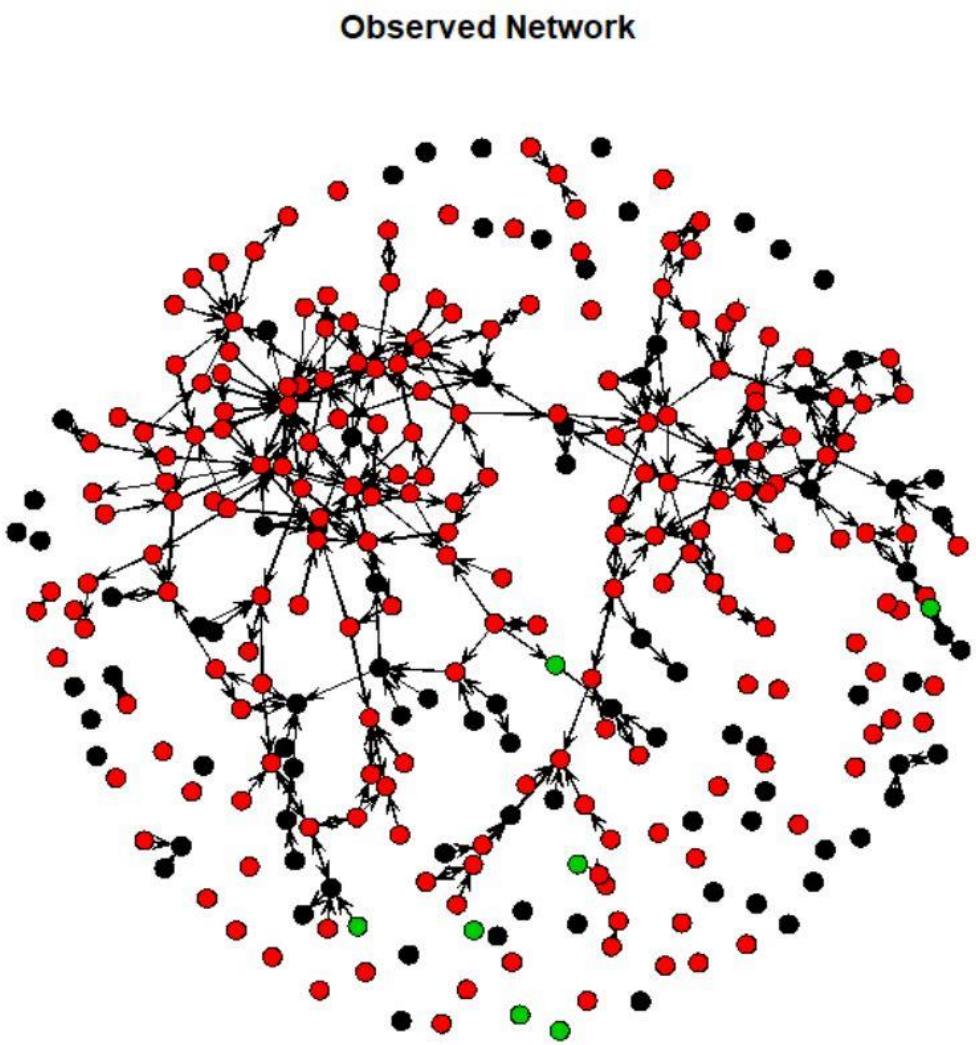
```
==========================
Summary of model fit
==========================

Formula:   net ~ edges + mutual + asymmetric("Gender", diff = FALSE) + asymmetric("Ethnicity",
    diff = FALSE) + nodematch("Ethnicity", diff = F) + nodefactor("Ethnicity") +
    simmelianties + isolates

Iterations:  9 out of 20

Monte Carlo MLE Results:
                           Estimate Std. Error MCMC %  p-value
edges                     -7.61993    0.24080      0  < 1e-04 ***
mutual                     7.01078    0.69272      0  < 1e-04 ***
asymmetric.Gender          0.94728    0.15397      0  < 1e-04 ***
asymmetric.Ethnicity       0.91375    0.34107      0 0.007385 **
nodematch.Ethnicity        1.49435    0.26764      1  < 1e-04 ***
nodefactor.Ethnicity.Latino -0.15842    0.03996      0  < 1e-04 ***
nodefactor.Ethnicity.White   0.19171    0.06642      0 0.003896 **
simmelianties              0.70382    0.18589      5 0.000153 ***
isolates                   1.20703    0.18657      0  < 1e-04 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

     Null Deviance: 118606  on 85556  degrees of freedom
 Residual Deviance:   3615  on 85547  degrees of freedom

AIC: 3633    BIC: 3717    (Smaller is better.)
> |
```

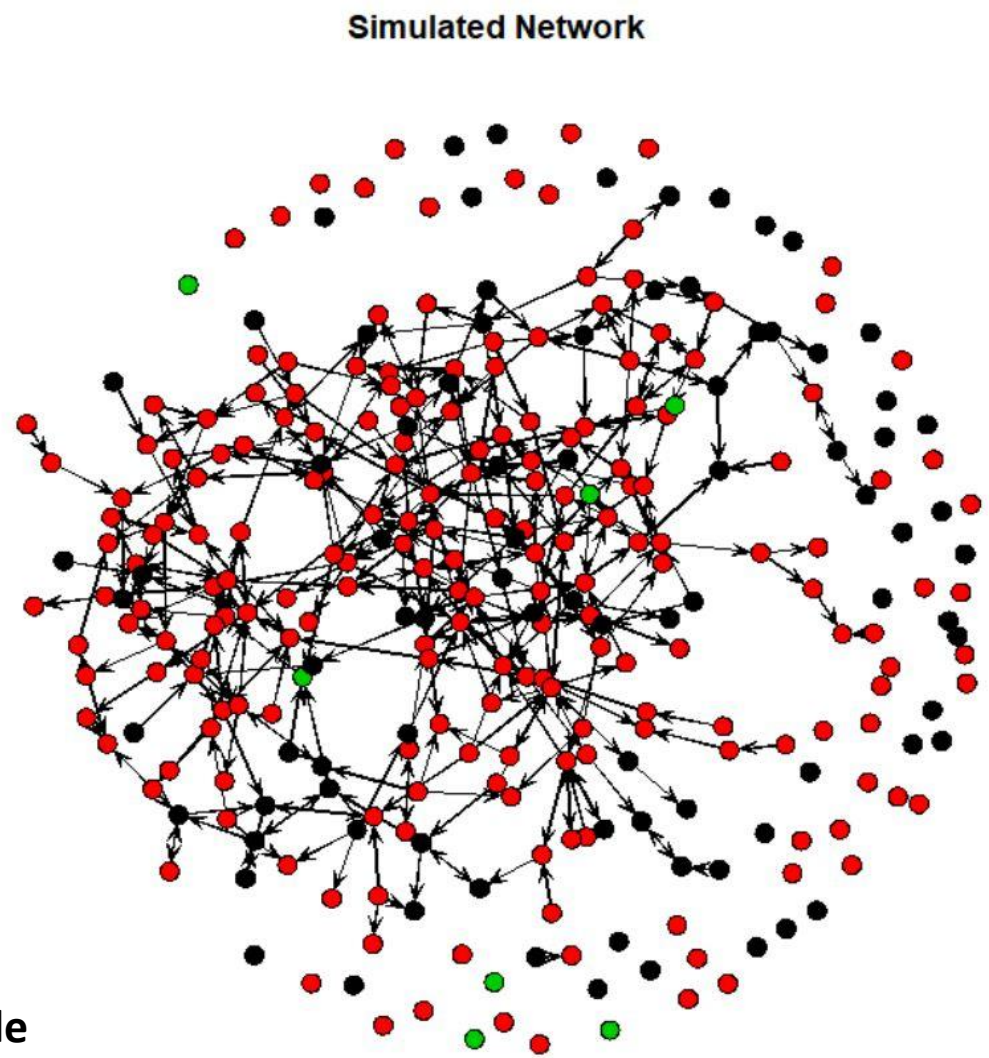
Model based on selected attributes





Model based on selected attributes





Red –Male

Black – Female

Green - Unknown

Conclusion

- The network is male-dominated and a high percentage of participants belongs from Latino ethnicity
- In-degree centrality is used to find important nodes and 12 nodes were selected to create the awareness about drugs as well as diseases
- Edges, mutuality, asymmetric dyads, isolates, homophily and ties in simmelian triads are the important local structures which represent some part of the global pattern in the network

References

1) WEEKS, M. R., CLAIR, S., BORGATTI, S. P., RADDA, K. & SCHENSUL, J. J. 2002. Social networks of drug users in high-risk sites: Finding the connections. AIDS and Behaviour, 6, 193-206.
2) Introduction to Exponential-family Random Graph (ERG or p*) modeling with ergm, The Statnet Devlopment Team, August 2017
3) ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks, Journal of Statistical Software, David R. Hunter et. al. (2008)
4) Speciation of Exponential-Family Random Graph Models: Terms and Computational Aspects, Journal of Statistical Software, Martina Morris et. al. (2008)
5) Social Network Analysis with sna, Journal of Statistical Software, Carter T. Butts (2008)
6) Exponential Random Graph (ERG or p*) Models, COMM 645: Communication Networks Annenberg School of Communication University of Southern California
7) A statnet Tutorial, Journal of Statistical Software, Steven M. Goodreau et. al. (2008)
8) https://rstudio-pubs-static.s3.amazonaws.com/157501_93a72a58ec614946901e10edf78c1384.html
9) https://www.sci.unich.it/~francesc/teaching/network/
10) http://badhessian.org/2012/09/lessons-on-exponential-random-graph-modeling-from-greys-anatomy-hook-ups/
11) http://www.mjdenny.com/Preparing_Network_Data_In_R.html