

## Appendix to: Understanding US counties: Clustering and Classification over voting, socio-economic and public health data

CHRISTINE E. DOIG-CARDET, Universitat Politècnica de Catalunya  
DIEGO GARCIA-OLANO, Universitat Politècnica de Catalunya

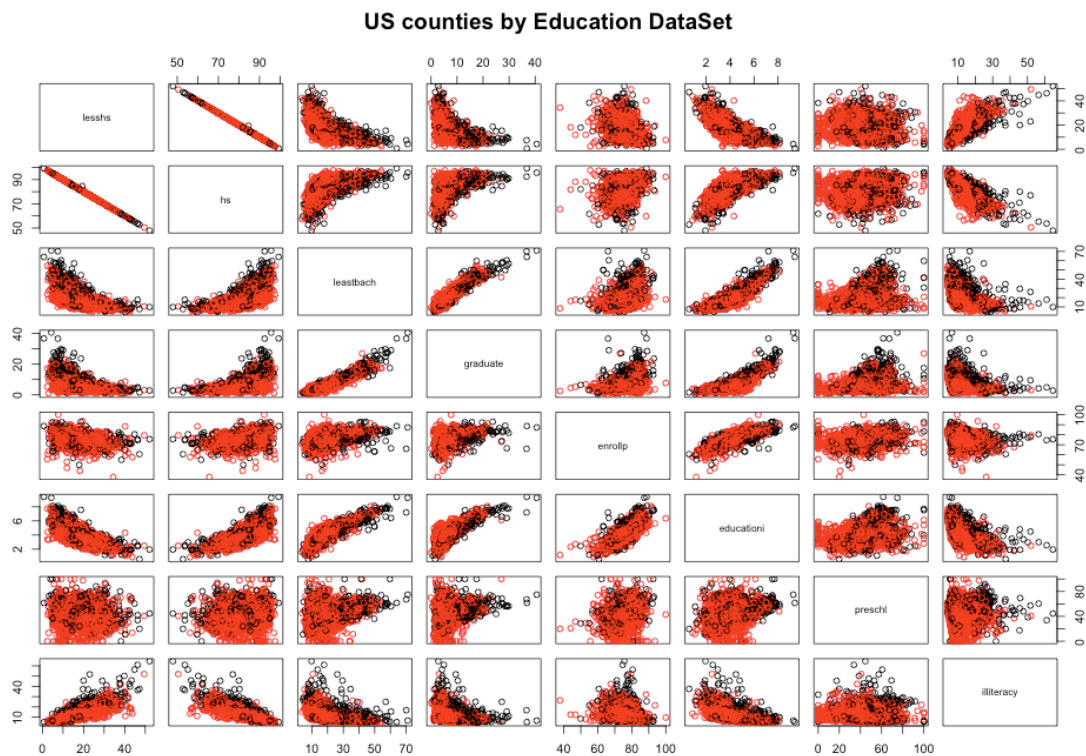
### A. VARIABLES DICTIONARY

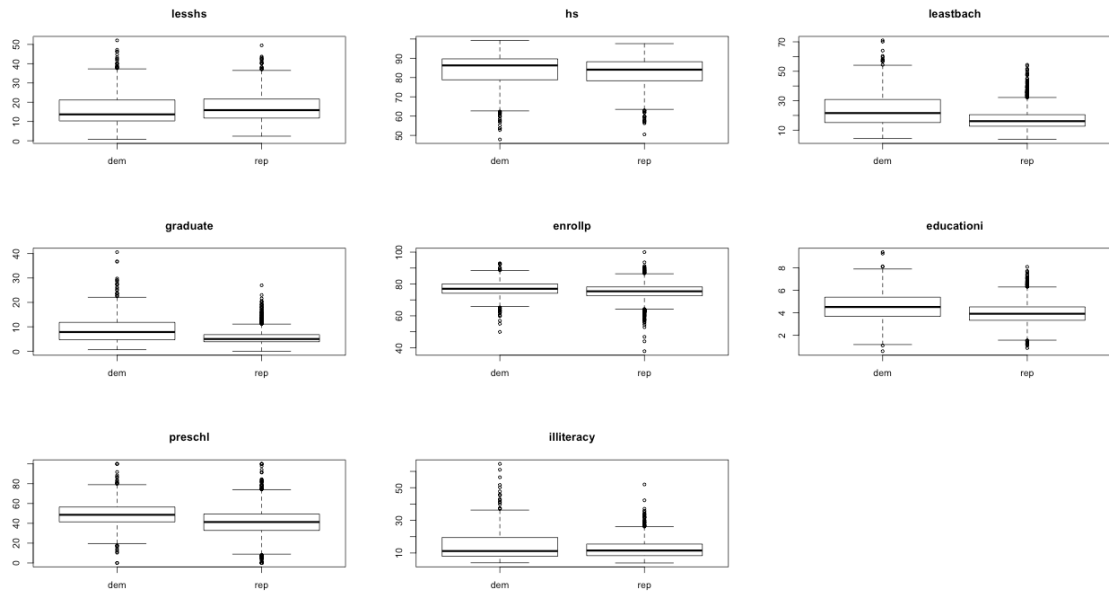
Our dataset consists in 3113 counties and 63 variables. The following table provides a full description of the variables.

VARIABLE	DESCRIPTION
<b>Education:</b>	
lesshs	% of population with less than high school degree
hs	% of population with high school degree
leastbach	% of population with bachelor's degree
graduate	% of population with graduate degree
enrollp	% of population between ages 3-24 enrolled in education system
educationi	education index
preschl	% of kids attending preschool
illiteracy	level of illiteracy among population
<b>Socio-economic:</b>	
earnings	earnings
incomei	income index
povo65	poverty of people over 65
belowpov	% population below poverty
gini	gini index
childpov.x	% of children in poverty
wmanage	% of workers in management positions
wsales	% of workers in sales
wfarm	% of workers in farming
wconstruct	% of workers in construction
wtransport	% of workers in transportation
unemployed	unemployed
Violentcrime	occurance of violent crimes
freelunch	%of free lunch in schools
<b>Health:</b>	
fairpoorhealth	% of populait
sickdays	days a year people take as sick days
mentaldays	days a year people take as mental days stress
smokers	smokers
adultobese	adults obesity
phyinactive	physically inactive
excsdrinking	excessive drinking
motordeath	deaths caused by motor accidents
stdsper100	std per 100.000 population
teenbirthrate	teen birth rate
uninsured.x	% of population uninsured
physicianratio	physician ratio
ambulatorycare	% of population in ambulatory care
diabeteshba1c	Percent of Diabetic Medicare enrollees receiving HbA1c test
mammography	mammography
noemotional support	No emotional support
refac	Access to recreational facilities. Recreational facility access rate per 100K population
limitedaccessheathyfood	Percent of people with limited access to healthy foods
fastfoodresp	Percent of restaurants that are fast food restaurants
diabectic	% of population who is diabetic (Centers for Disease Control (CDC), Small Area Obesity Estimates)
mphratio	%mental health providers
healthcosts	health costs

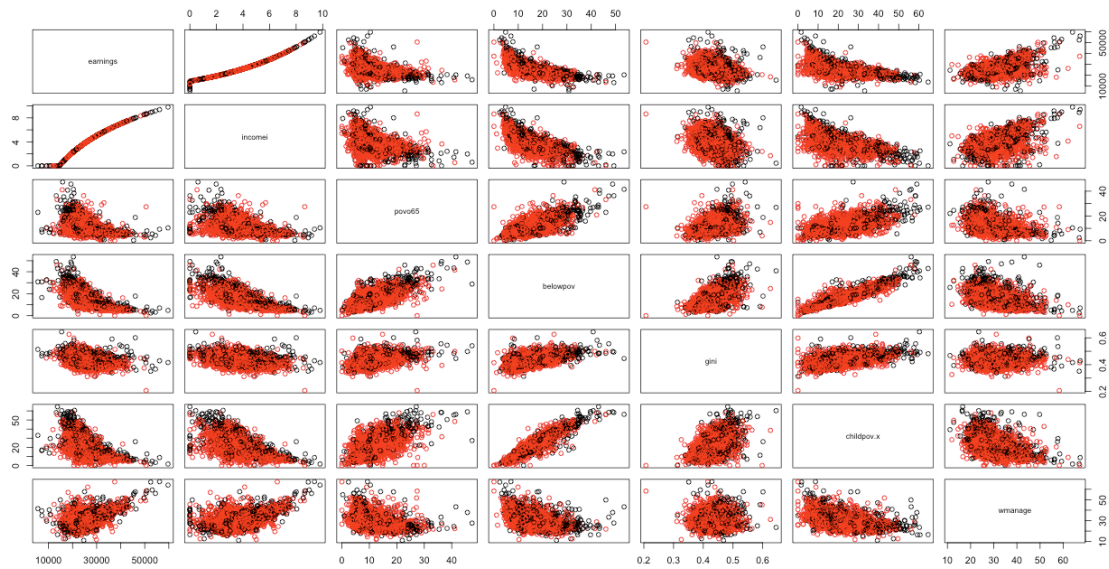
healthyfood	Percent of zip codes in county with healthy food outlets ( according to Census Zip Code Business Patterns). Healthy food outlets include grocery stores (NAICS 445110) with > 4 employees and produce stands/farmers' markets (NAICS 445230)
dentistratio	% of dentists
<b>Demographics:</b>	
white	% of whites
afric	% of African-american
nativ	% of narive american
asian	% of asian
other	% of other races
latin	% of latinos
popul	population
lowbirthrate	birth rate
singleparent	single parents
under18	%population under 18
over65	% population over 65
noenglish	% population no english
female	% population female
rural	% rural
drivealone	Commuting alone to work. % of people who commute to work alone

## B. PAIRS PLOT OF VARIABLES BY DEMOCRAT/REPUBLICAN WINNER

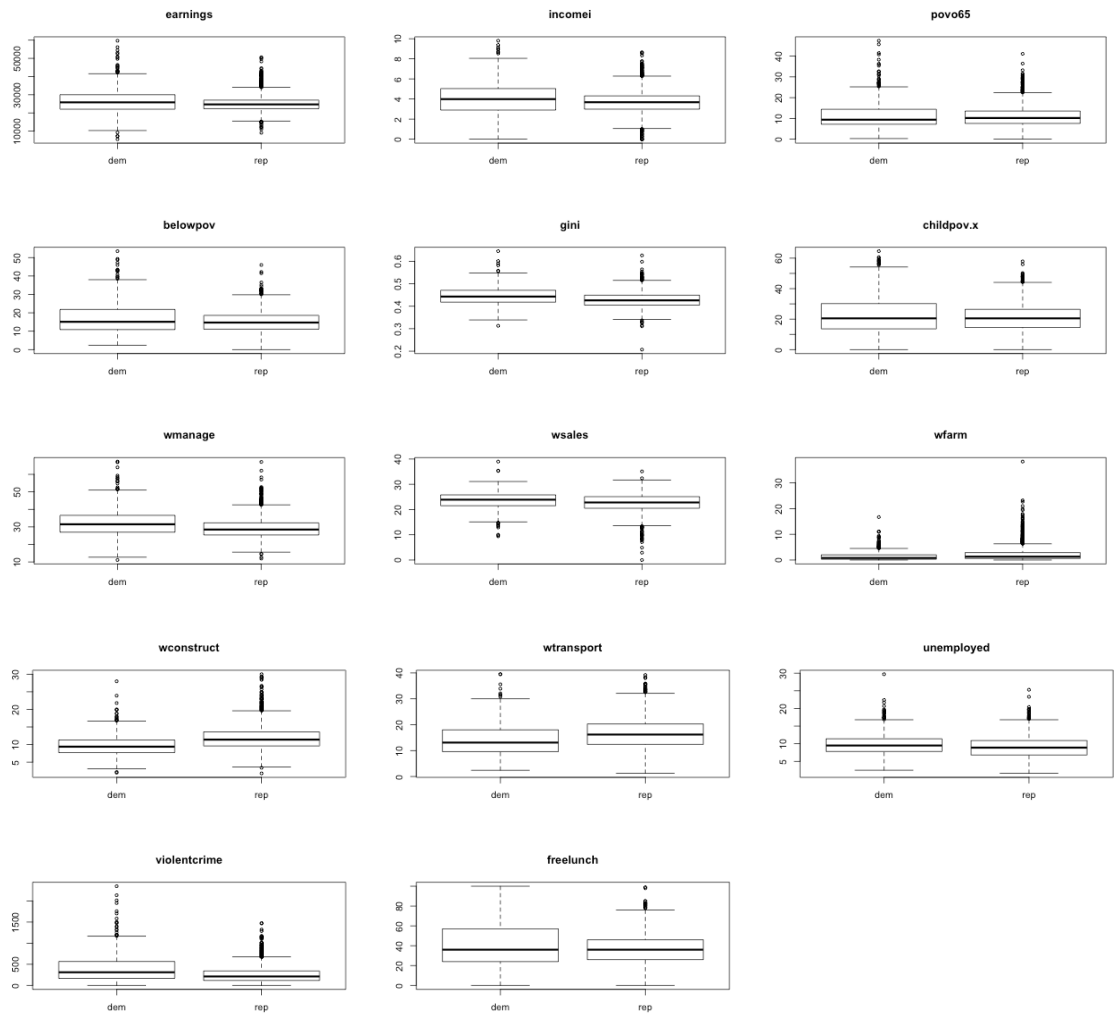
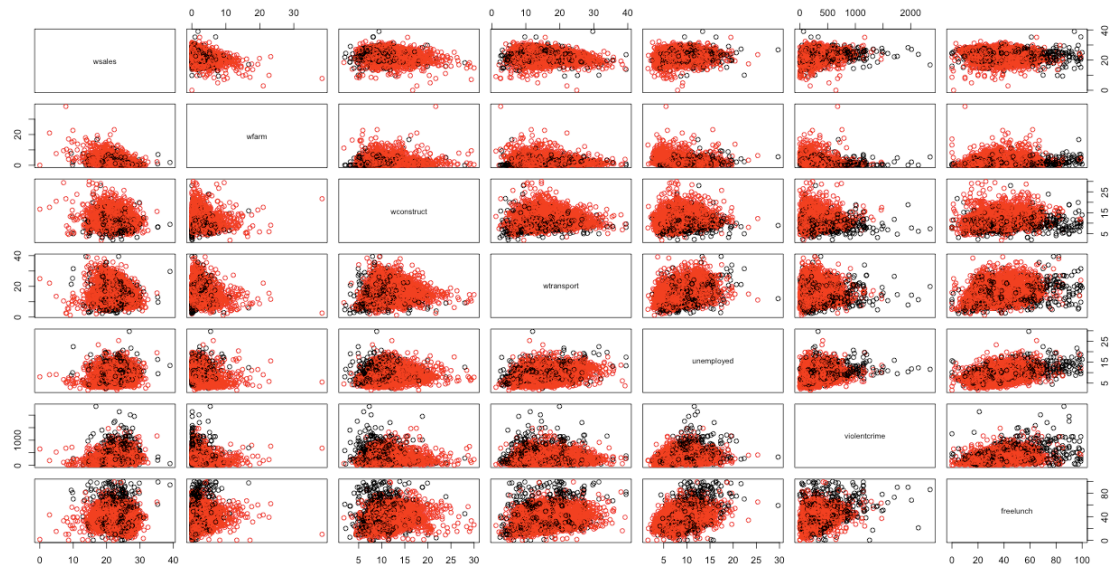




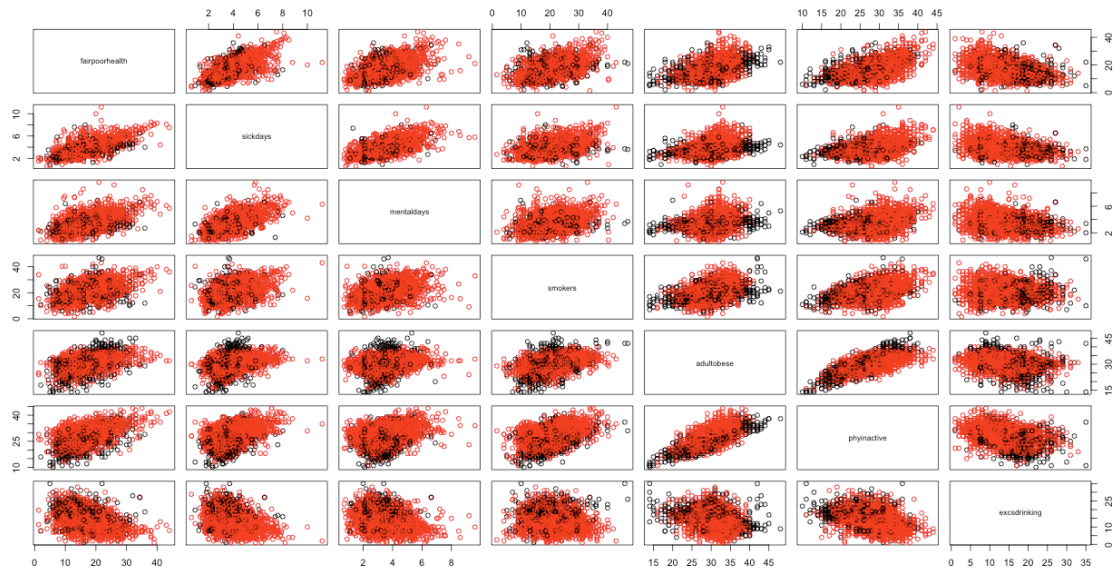
US counties Socio-Economic Variables by Party (I)



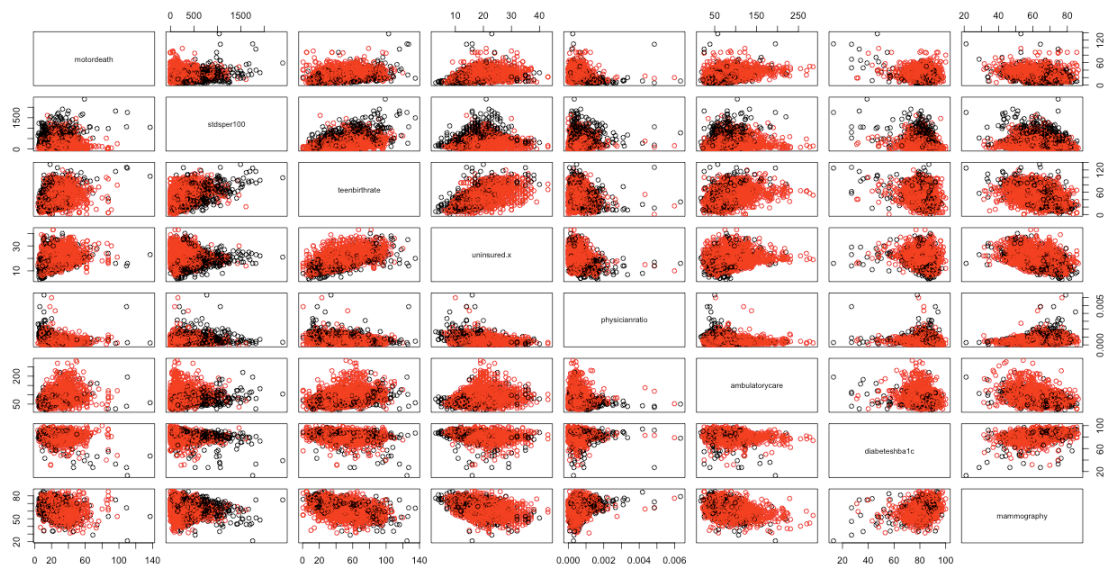
# US counties Socio-Economic Variables by Party (II)



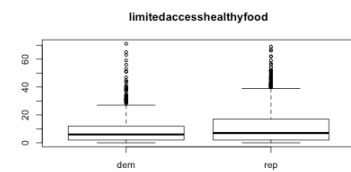
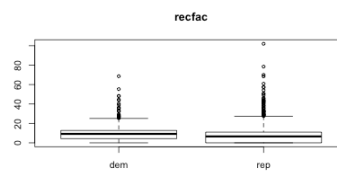
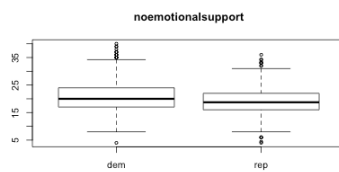
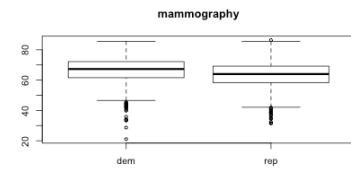
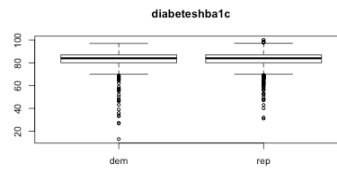
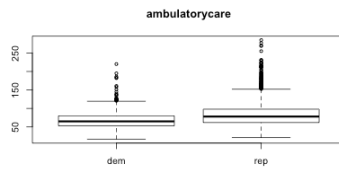
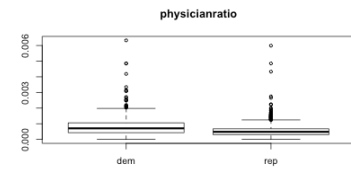
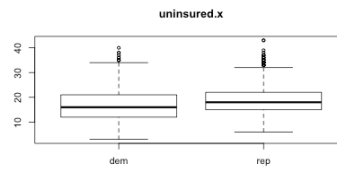
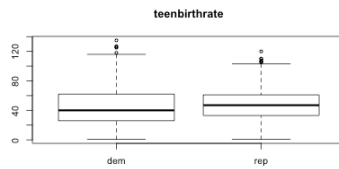
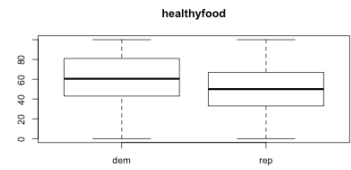
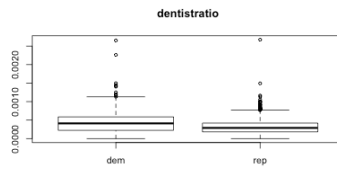
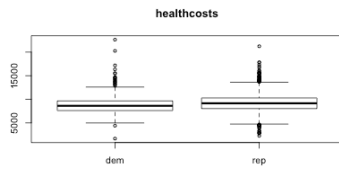
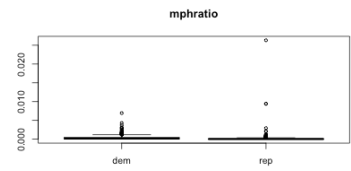
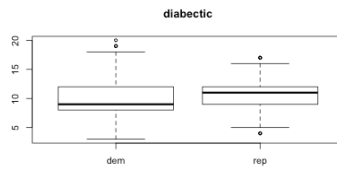
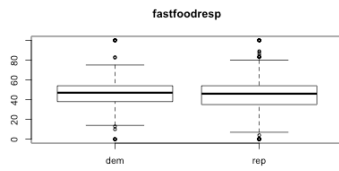
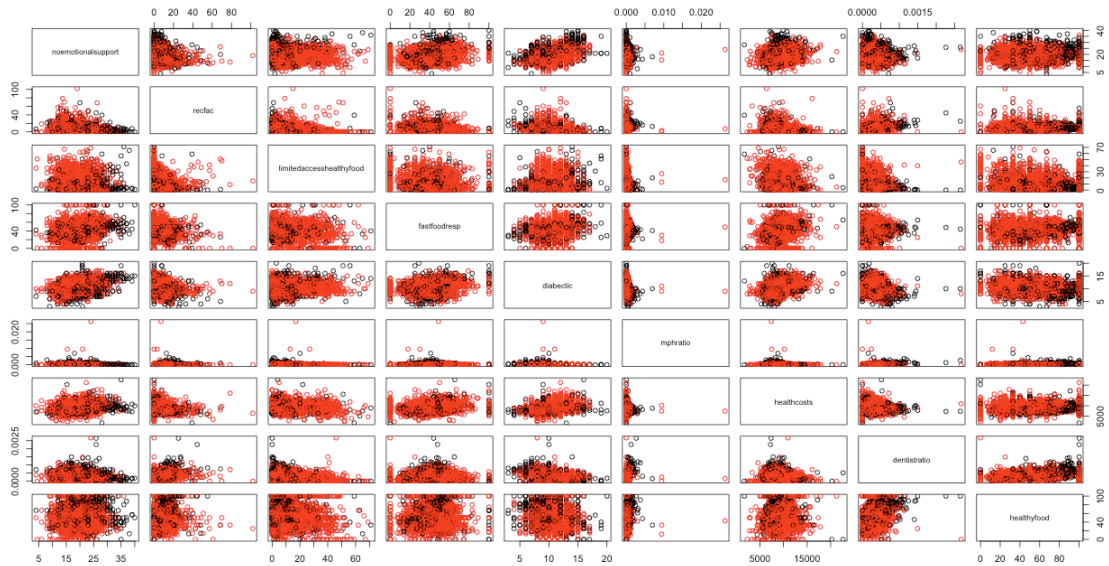
US counties Health Variables by Party (I)



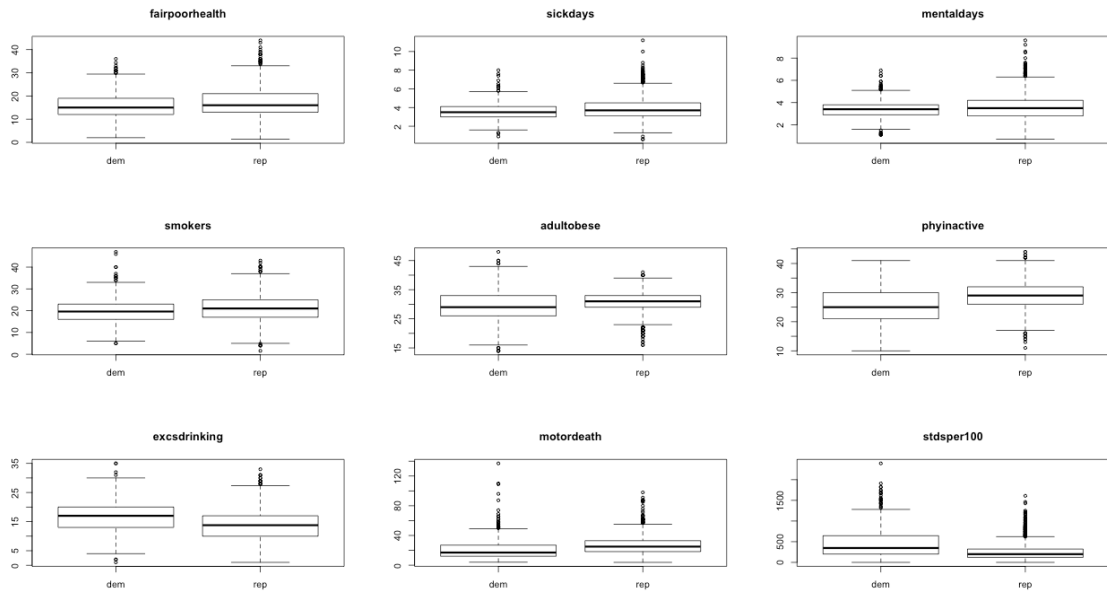
US counties Health Variables by Party (II)



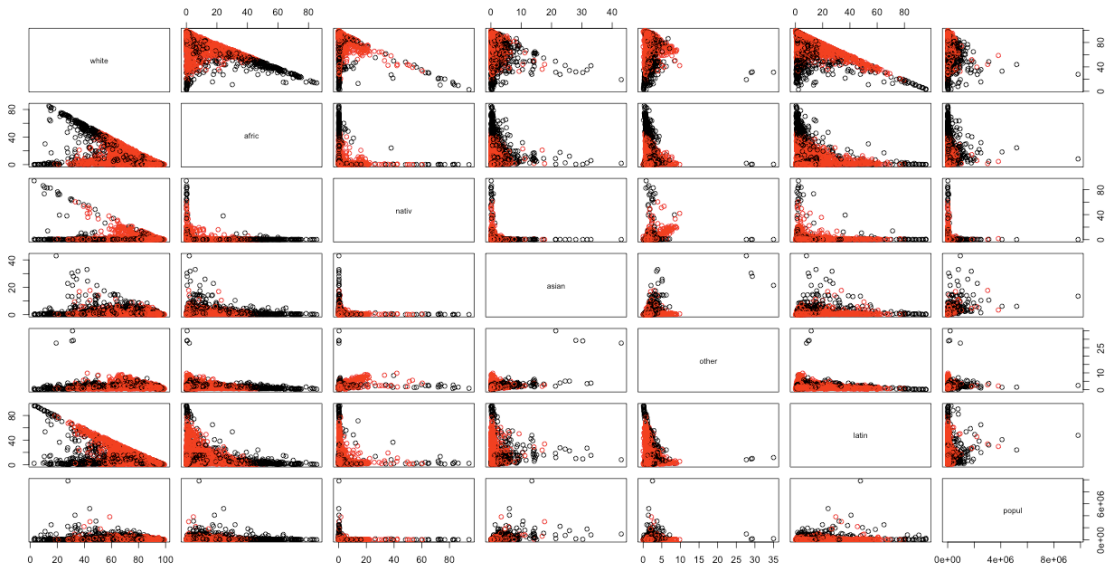
US counties Health Variables by Party (III)



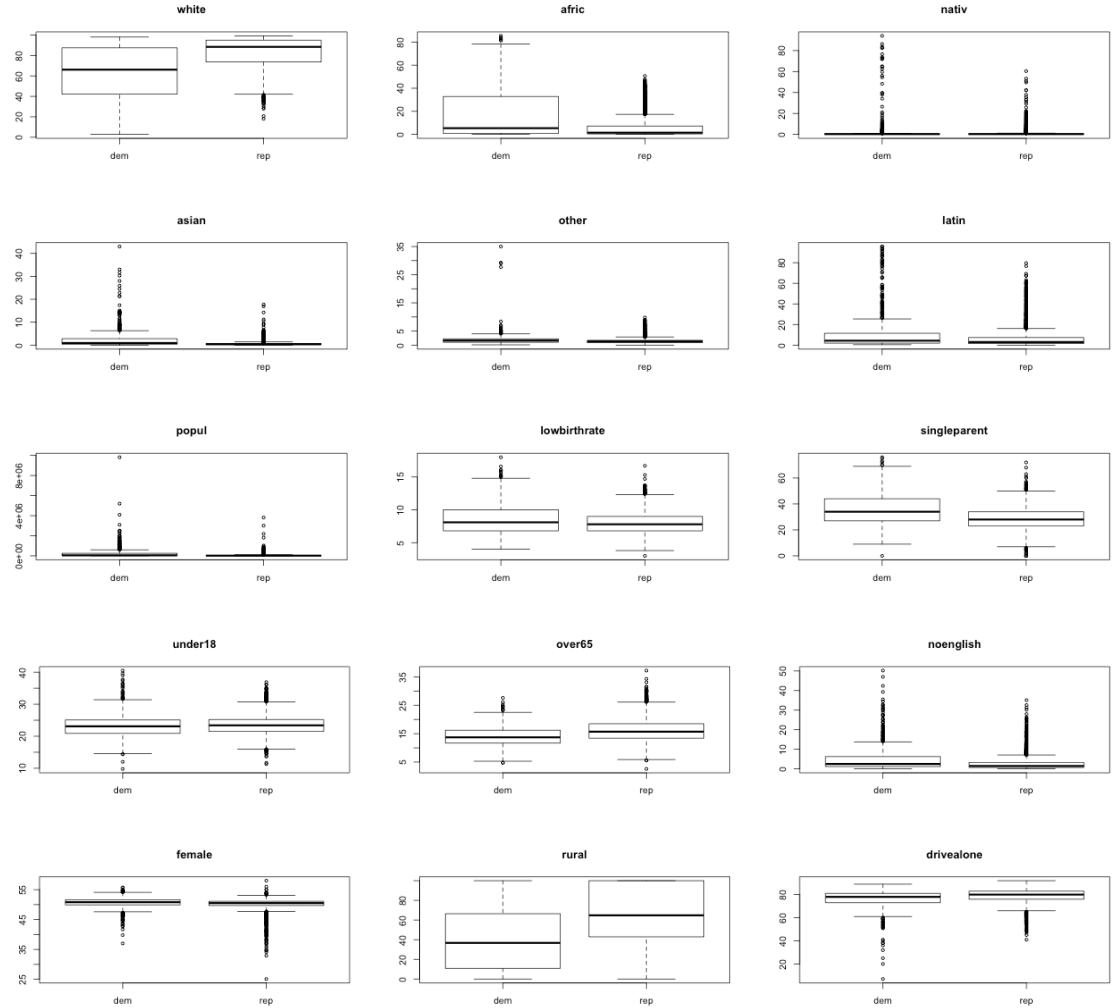
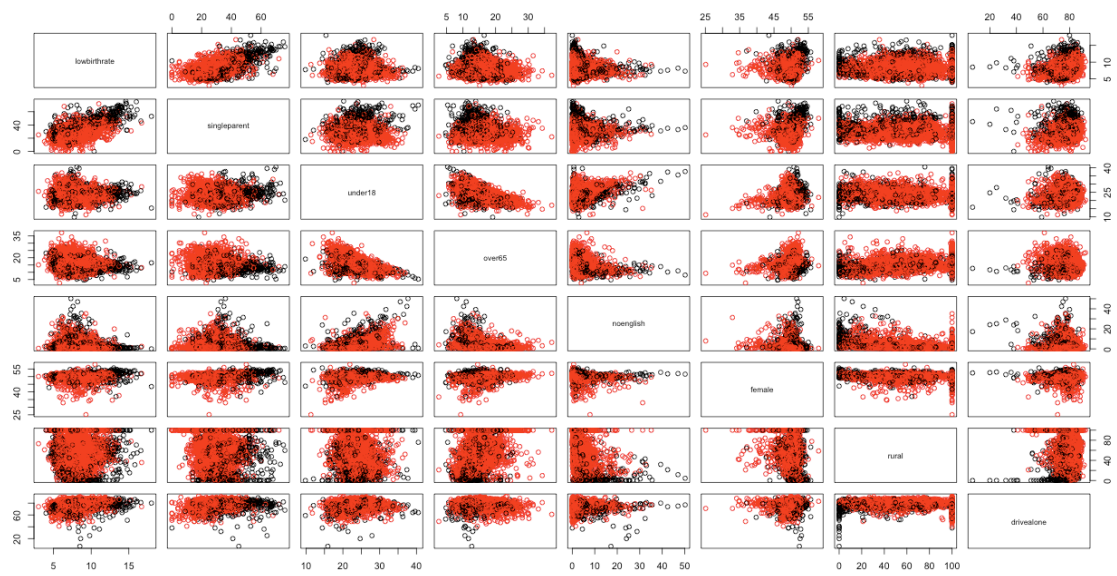




US counties Demographics Variables by Party (I)



US counties Demographics Variables by Party (II)





---

---