MIS581 Portfolio Project

Hailey Tibbetts

1/4/2021

#Import Dataset

Nutrition\_\_Physical\_Activity\_\_and\_Obesity\_.\_Women\_\_Infant\_\_and\_Child <- read.csv("~/Desktop/Classwork/Nutrition\_\_Physical\_Activity\_\_and\_Obesity\_-\_Women\_\_Infant\_\_and\_Child.csv") View(Nutrition\_\_Physical\_Activity\_*and\_Obesity*.\_Women\_\_Infant\_\_and\_Child) dim(Nutrition\_\_Physical\_Activity\_\_and\_Obesity\_.\_Women\_\_Infant\_\_and\_Child) df<-Nutrition\_\_Physical\_Activity\_\_and\_Obesity\_.\_Women\_\_Infant\_\_and\_Child install.packages(“PANDA”) full<-df

#Subset by category

new<-data.frame(fullQuestion, fullfull.Question==‘Percent of WIC children aged 2 to 4 years who have an overweight classification’) subset(new, newfull.Question==‘Percent of WIC children aged 2 to 4 years who have an overweight classification’) View(overweight) infant<-subset(new, newfull.Question==‘Percent of WIC children aged 2 to 4 years who have obesity’) rowsum(overweight$full.LocationAbbr==‘AK’) View(obesity)

#Tables table(fullQuestion) locationvsquestion<-table(fullQuestion) locationvsrace<-table(fullRace.Ethnicity) locationvsrace valuevsstate<-table(fullData\_Value) valuevsstate barplot(table(fullLocationAbbr, fullData\_Value)

#Separating by state

ak<-subset(obesity, obesityfull.LocationAbbr==‘AK’) aL<-subset(obesity, obesityfull.LocationAbbr==‘AR’) az<-subset(obesity, obesityfull.LocationAbbr==‘CA’) co<-subset(obesity, obesityfull.LocationAbbr==‘CT’) dc<-subset(obesity, obesityfull.LocationAbbr==‘DE’) fl<-subset(obesity, obesityfull.LocationAbbr==‘GA’) gu<-subset(obesity, obesityfull.LocationAbbr==‘HI’) ia<-subset(obesity, obesityfull.LocationAbbr==‘ID’) il<-subset(obesity, obesityfull.LocationAbbr==‘IN’) indi<-subset(obesity, obesityfull.LocationAbbr==‘KS’) ky<-subset(obesity, obesityfull.LocationAbbr==‘LA’) ma<-subset(obesity, obesityfull.LocationAbbr==‘MD’) me<-subset(obesity, obesityfull.LocationAbbr==‘MI’) mn<-subset(obesity, obesityfull.LocationAbbr==‘MO’) ms<-subset(obesity, obesityfull.LocationAbbr==‘MT’) nc<-subset(obesity, obesityfull.LocationAbbr==‘ND’) ne<-subset(obesity, obesityfull.LocationAbbr==‘NH’) nj<-subset(obesity, obesityfull.LocationAbbr==‘NM’) nv<-subset(obesity, obesityfull.LocationAbbr==‘NY’) oh<-subset(obesity, obesityfull.LocationAbbr==‘OK’) or<-subset(obesity, obesityfull.LocationAbbr==‘PA’) pr<-subset(obesity, obesityfull.LocationAbbr==‘RI’) sc<-subset(obesity, obesityfull.LocationAbbr==‘SD’) tn<-subset(obesity, obesityfull.LocationAbbr==‘TX’) ut<-subset(obesity, obesityfull.LocationAbbr==‘VA’) vi<-subset(obesity, obesityfull.LocationAbbr==‘VT’) wa<-subset(obesity, obesityfull.LocationAbbr==‘WI’) wv<-subset(obesity, obesityfull.LocationAbbr==‘WY’)

#Removing missing values wy2<-wy[-c(45),] View(wy2) dc2<-dc[-c(14,19,32,40,45),] View(dc2) vt2<-vt[-c(4),] tn2<-tn[-c(3,9, 25),] View(tn2) nh2<-nh[-c(7, 21, 31),] View(ms) ms2<-ms[-c(49, 32, 26),] View(ms2) hi2<-hi[-c(45, 26),] View(hi2) de2<-de[-c(35, 30, 23, 20, 19),] View(de2) ct2<-ct[-c(5),] View(ct2)

#Averages by state akavg<-sum(akfull.Data\_Value)/55 aLavg<-sum(aLfull.Data\_Value)/55 azavg<-sum(azfull.Data\_Value)/55 coavg<-sum(cofull.Data\_Value)/55 dcavg<-sum(dcfull.Data\_Value)/55 flavg<-sum(flfull.Data\_Value)/55 guavg<-sum(gufull.Data\_Value)/55 iaavg<-sum(iafull.Data\_Value)/55 ilavg<-sum(ilfull.Data\_Value)/55 meavg<-sum(mefull.Data\_Value)/55 maavg<-sum(mafull.Data\_Value)/55 kyavg<-sum(kyfull.Data\_Value)/55 miavg<-sum(mifull.Data\_Value)/55 moavg<-sum(mofull.Data\_Value)/55 mtavg<-sum(mtfull.Data\_Value)/55 ndavg<-sum(ndfull.Data\_Value)/55 ndavg<-sum(ndfull.Data\_Value)/55 nhavg<-sum(nhfull.Data\_Value)/55 nmavg<-sum(nmfull.Data\_Value)/55 nyavg<-sum(nyfull.Data\_Value)/55 okavg<-sum(okfull.Data\_Value)/55 paavg<-sum(pafull.Data\_Value)/55 riavg<-sum(rifull.Data\_Value)/55 sdavg<-sum(sdfull.Data\_Value)/55 txavg<-sum(txfull.Data\_Value)/55 utavg<-sum(vafull.Data\_Value)/55 vaavg<-sum(vafull.Data\_Value)/55 vtavg<-sum(vtfull.Data\_Value)/55 wiavg<-sum(wifull.Data\_Value)/55 wyavg<-sum(wyfull.Data\_Value)/54 dcavg<-sum(dc2full.Data\_Value)/50 hiavg<-sum(hi2full.Data\_Value)/52 nhavg<-sum(nh2full.Data\_Value)/52 wyavg<-sum(wy2full.Data\_Value)/54

#Separating states by region southavg<-sum(txavg,okavg,aravg,laavg,msavg,aLavg,tnavg,kyavg,wvavg,mdavg,deavg,vaavg,ncavg,scavg,gaavg,flavg)/16 northeastavg<-sum(meavg,nhavg,vtavg,nyavg,paavg,njavg,ctavg,riavg,maavg)/9 southeastavg<-sum(wvavg,kyavg,vaavg,tnavg,ncavg,scavg,gaavg,aLavg,msavg,aravg,laavg,flavg)/12 southwestavg<-sum(txavg,okavg,nmavg,azavg)/4 westavg2<-sum(coavg,wyavg,mtavg,idavg,waavg,oravg,utavg,caavg,nvavg,akavg,hiavg)/11northeastavg<-sum(meavg,maavg,riavg,ctavg,nhavg,vtavg,nyavg,paavg,njavg,deavg,mdavg)/11 midwestavg<-sum(ohavg,indiavg,miavg,ilavg,moavg,wiavg,mnavg,iaavg,ksavg,neavg,sdavg,ndavg)/12

#Pie chart per region regions<- c(westavg2, southeastavg, southwestavg, northeastavg, midwestavg) barplot(regions) lables<- c(“West”, “Southeast”, “Southwest”, “Northeast”, “Midwest”) pie(regions, labels = lables, main=“Percentages per Region”)