

Zhengyu Pan
46020020
Assignment 2 Task 3
May 13th, 2020

First of all here is my github link

https://github.com/zacharyPan/zhengyuPanBUSA8090_A2T3

Please check code download from github instead of copy and paste from pdf.

Question 1

Part a,

The author tried to combine `mi_counties` into `mi_census` (left join) by using “country” from `mi_census` equal to “id” from `mi_counties`. That way, author could access all the information with countries in `census_counties`.

Part b,

The author tried to connect map with wdi database, using country code. that way, the author could showed all wdi information on a map.

Question 2, (question2.R on github)

The code will not work unless all the libraries were installed. Additionally, install “UserNetR” by

`install_github("DougLuke/UserNetR")`

Citation,

(As_tibble Function | R Documentation, n.d.) *(First Line of Every R Script?, 2017)* *(“Select & Rename R Functions of Dplyr Package (2 Examples),” n.d.)*

```
# Assignment 2 Task 3  
# BUSA8090  
# Author, Zhengyu Pan  
# May 13th, 2020
```

```
rm(list = ls(all = TRUE)) # clear all  
# additionally do shift + control+ F10  
# to clean up the libraries
```

```

library(WDI) # assume that WDI is installed
library(ggplot2) # assume that ggplot2 is installed
library(dplyr) # assume that dplyr is installed
library(maps) # assume that maps is installed
               # it seems that we don't need this library, but
               # I will leave it here anyway
library(countrycode) # assume that countrycode is installed
library(UserNetR)    # assume already did
                     # install_github("DougLuke/UserNetR")
                     # install.packages("UserNetR") will NOT work
library(sna)         # assume sna is installed

# question 2 part a
wdi_data <- WDI(indicator=c("NY.GDP.PCAP.PP.KD", "EN.ATM.CO2E.PC"), start =
2010, end = 2010, extra = TRUE)
wdi_data = as_tibble(wdi_data)
wdi_data <- wdi_data %>% rename(GDPpercap = NY.GDP.PCAP.PP.KD,
Emit_co2percap=EN.ATM.CO2E.PC)
dat_map<-map_data("world")
dat_map$ccode <- countrycode(dat_map$region,origin = "country.name",destination =
"wb")
wdi$ccode<- countrycode(wdi$country, origin = "country.name", destination = "wb")
merged <-full_join(dat_map, wdi, by="ccode")
ggplot(merged, aes(x=long,y =lat, group = group, fill= Emit_co2percap))+geom_polygon()

# question 2 part b

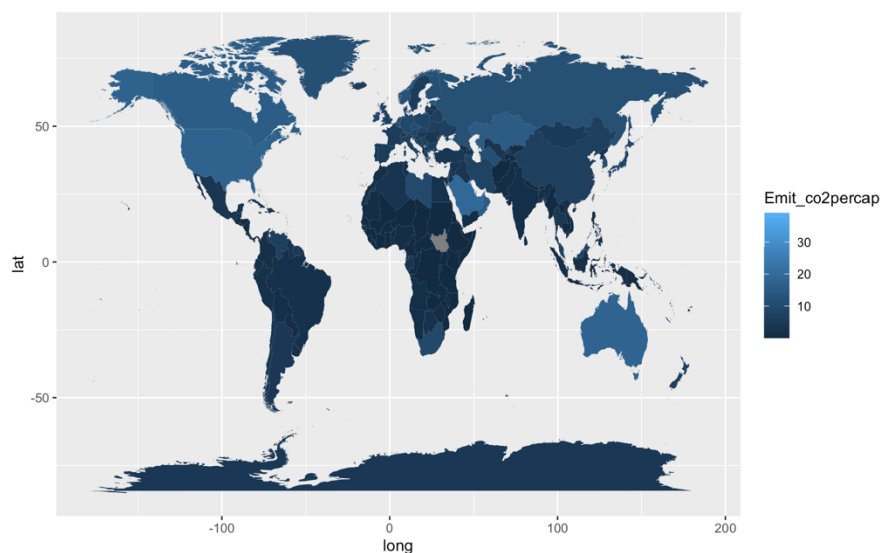
data(Bali)
op <- par(mar = c(0,0,0,0))
plot(Bali,displaylabels=TRUE,label.cex=0.8,pad=0.4,label.col="darkblue")

# question 2 part c

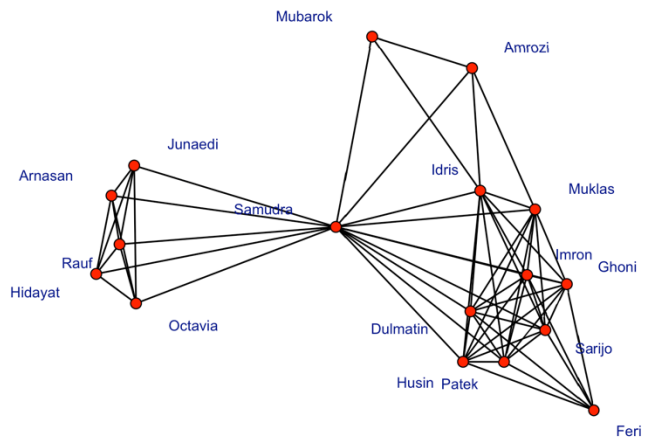
rolelab <- get.vertex.attribute(Bali,"role")
plot(Bali,usearrows=FALSE,label=rolelab,displaylabels=T,label.col="darkblue")

```

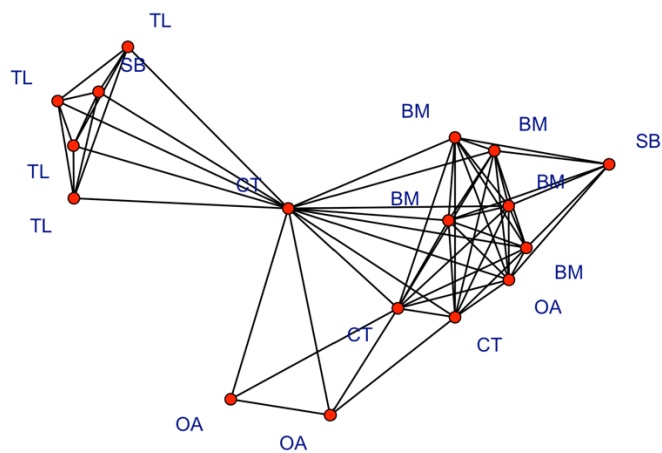
Part a



Part b



Part c



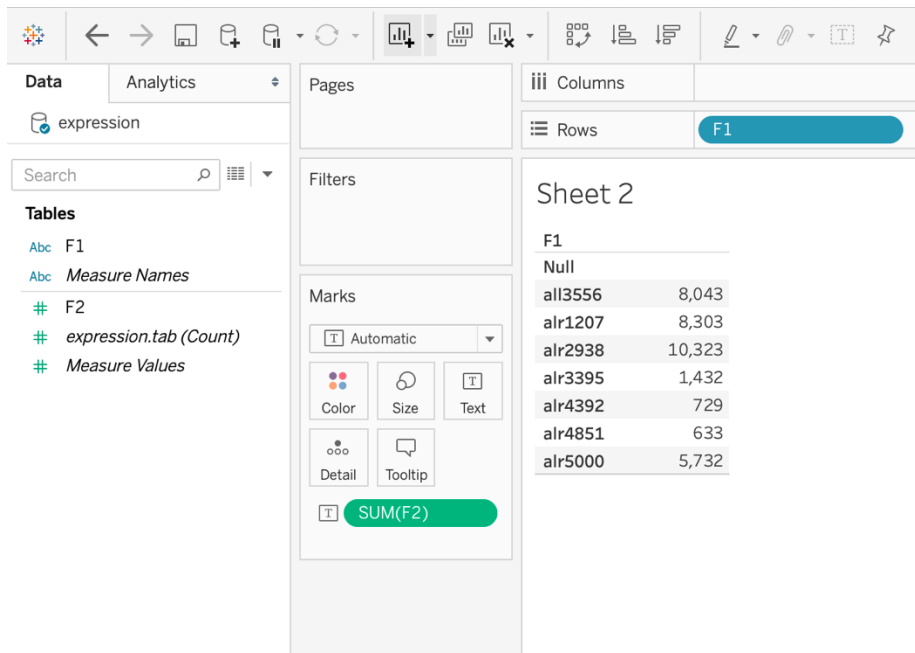
Question 3 (question3a.twb question3b.twb question3c.twb)

Please use the data I upload on github to check, since annotation.tab I need manually add header to make it work

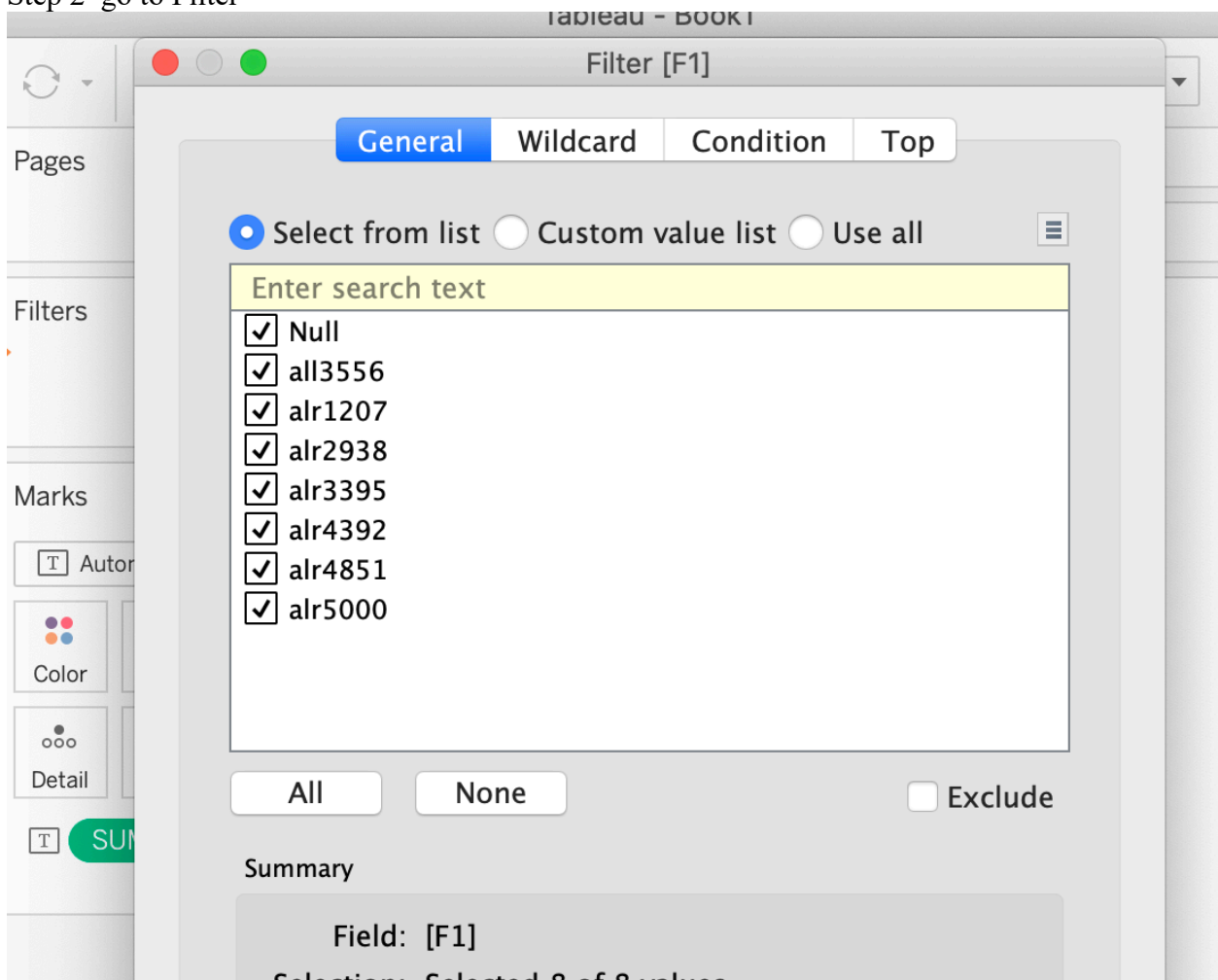
Citation,
(Using “CONTAINS” Function to Filter for Certain... |Tableau Community Forums, n.d.)

Part a,

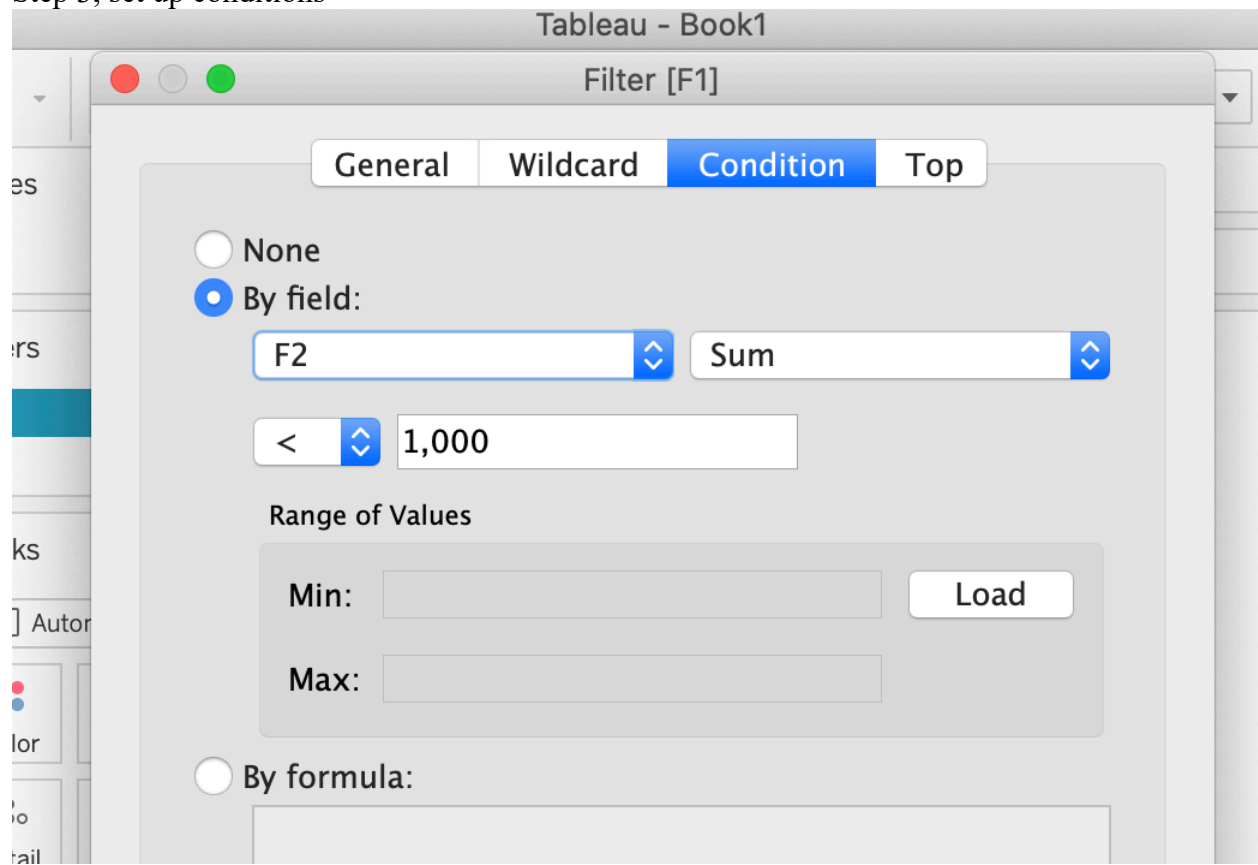
Step 1,



Step 2 go to Filter



Step 3, set up conditions



Step 4 result

Columns

Rows

F1

Filters

F1

Marks

Automatic

Color

Size

Text

Detail

Tooltip

SUM(F2)

Sheet 2

F1

alr4392	729
alr4851	633

Part b,

Step 1, I need add a header manually, so it won't use any real data as header

Filter [C2]

General Wildcard **Condition** Top

☐ None

☐ By field:

C1 Count

= 0

Range of Values

Min: Load

Max:

☒ By formula:

`CONTAINS([C2] , "ase")`

Step 4

Pages

Columns

Rows

C1

C2

C3

Filters

C2

Marks

Aut...

Co...

Size

Text

De...

To...

Sheet 1

C1	C2	C3	
alr3395	adenylosuccinate lyase	Purine biosynthesis	Abc
alr5000	CTP synthetase	Pyrimidine biosynthesis	Abc
alr2938	iron superoxide dismutase	Detoxification	Abc
alr4851	preprotein translocase su..	Protein and peptide secre..	Abc
all3556	succinate-dehydrogenase	TCA cycle	Abc
alr1207	uridylate kinase	Pyrimidine biosynthesis	Abc

Part 3,

Step 1 after load data like part 2

C3

Filter...

Show Filter

Show Highlighter

Sort...

Format...

✓ Show Header

e biosynthesis

Step 2

General

Wildcard

Condition

Top

☐ None

☐ By field:

C1



Count



=



0

Range of Values

Min:

Load

Max:

☒ By formula:

```
CONTAINS([C3], "Pyrimidine biosynthesis")  
OR CONTAINS([C3], "Purine biosynthesis")
```

Step 3



C1	C2	C3	
alr1207	uridylate kinase	Pyrimidine biosynthesis	Abc
alr3395	adenylosuccinate lyase	Purine biosynthesis	Abc
alr5000	CTP synthetase	Pyrimidine biosynthesis	Abc

Reference:

As_tibble function | *R Documentation*. (n.d.). Retrieved May 12, 2020, from

https://www.rdocumentation.org/packages/tidyquant/versions/0.5.3/topics/as_tibble

First line of every R script? (2017, September 17). RStudio Community.

<https://community.rstudio.com/t/first-line-of-every-r-script/799>

Select & rename R Functions of dplyr Package (2 Examples). (n.d.). *Statistics Globe*. Retrieved

May 12, 2020, from <https://statisticsglobe.com/r-select-rename-functions-dplyr-package/>

Using “CONTAINS” function to filter for certain... | *Tableau Community Forums*. (n.d.).

Retrieved May 12, 2020, from <https://community.tableau.com/thread/122402>