PROJECT\_A

EDGAR

5/17/2022

library(tidyverse) library(janitor) library(ggridges) library(patchwork)

# import data

alc <- read\_csv (“alcohol-consumption-vs-gdp-per-capita.csv”)

# clean data and rename vars

alc <- alc %>% clean\_names() %>% # clean up column names mutate(alcohol\_consumption = total\_alcohol\_consumption\_per\_capita\_liters\_of\_pure\_alcohol\_projected\_estimates\_15\_years\_of\_age) %>% # shortening column name mutate(gdp\_per\_cap = gdp\_per\_capita\_ppp\_constant\_2017\_international) %>% # shortening column name mutate(population\_est = population\_historical\_estimates) %>% # shortening column name select(entity, year, alcohol\_consumption, gdp\_per\_cap, population\_est, continent) %>% # selecting wanted columns drop\_na() %>% # remove n/a’s filter(year == 2015) # filter for year - 2015

# view data

head(alc)

# visualize / explore data

# alc consumption plot

p1 <- alc %>% ggplot(aes(x = alcohol\_consumption, y = continent, fill = continent)) + geom\_density\_ridges(aes(alpha = .03)) + theme(legend.position = “none”)

# gdp plot

p2 <- alc %>% ggplot(aes(x = gdp\_per\_cap, y = continent, fill = continent)) + geom\_density\_ridges(aes(alpha = .03)) + theme(legend.position = “none”)

# population plot

p3 <- alc %>% ggplot(aes(x = population\_est, y = continent, fill = continent)) + geom\_density\_ridges(aes(alpha = .03)) + theme(legend.position = “none”) p1 / p2 / p3

alc %>% ggplot(aes(x = log(population\_est), y = gdp\_per\_cap)) + geom\_point(aes(color = alcohol\_consumption)) + facet\_wrap(~continent)

#analysis alc <- alc %>% mutate(alcohol\_use = case\_when( alcohol\_consumption < 5 ~ “low”, alcohol\_consumption < 10 ~ “medium”, alcohol\_consumption > 12 ~ “high” ))

head(alc)

alc %>% drop\_na() %>% ggplot(aes(x = log(population\_est), y = gdp\_per\_cap)) + geom\_jitter(aes(color = alcohol\_use)) + facet\_wrap(~alcohol\_use, nrow = 2)

# which countries in 2015 had high alcohol use?

alc %>% filter(alcohol\_use == “high”) %>% select(entity, continent)

# which countries in 2015 had medium alcohol use?

alc %>% filter(alcohol\_use == “medium”) %>% select(entity, continent)

# what countries had low alcohol use?

alc %>% filter(alcohol\_use == “low”) %>% select(entity, continent)