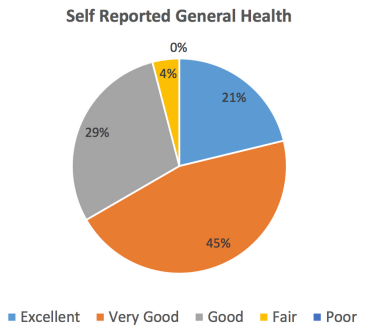
Brian Lambert
STA 404
Module 5 Homework

1. Self-Reported General Health Pie Chart Analysis

a. perspective on difficulty of executing perceptual tasks



According to Cleveland and McGill, the primary elementary task of a pie chart is to extract numeric information based on the perception of angles, but pie charts are tricky because the areas and arc lengths are variable. Given this information, I believe a pie chart is an acceptable display for this data because of the simplicity of noticing the differences in angles between each of the slices. It is easy to notice the difference in angles because there are only 4 slices and each has a rather distinct angle. If there were more slices or slices had similar angles, then a pie chart would likely not be an ideal visualization.

b. use of color in this pie chart

Knaflic states that color should be used to direct the audience's attention, and that a color pallet should should be grey scale with pointed use of color. I would say that this pie chart does use color to attract the audience's attention to the chart itself, however, the color scheme would disappoint Knaflic. It uses a seemingly random, non-grey scale theme that does not give the audience more intuition about the purpose of the graphic. I would likely use a scheme that colors the poor as red and then follows a hue to excellent as green. I like this scheme because the slices are ordinal and should be colored accordingly. Another way to show this ordinal data that Knaflic would probably like is to use a grey scale that increases or decreases in darkness from poor to excellent.

c. modification or alternative

As Cleveland and McGill suggest, I would like to display this pie chart data as a divided bar chart with each bar corresponding to a slice. I believe this is a better display because it is easier to compare the height of bars than the angles of a pie chart. This change would make it easier for the audience to get the overall story of that data and lessen the chance of someone misinterpreting the pie chart.

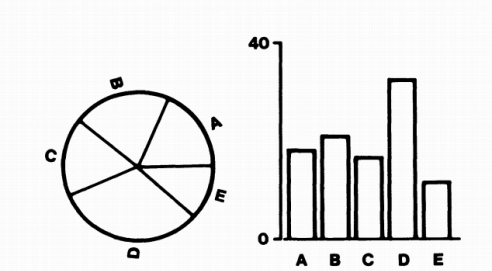
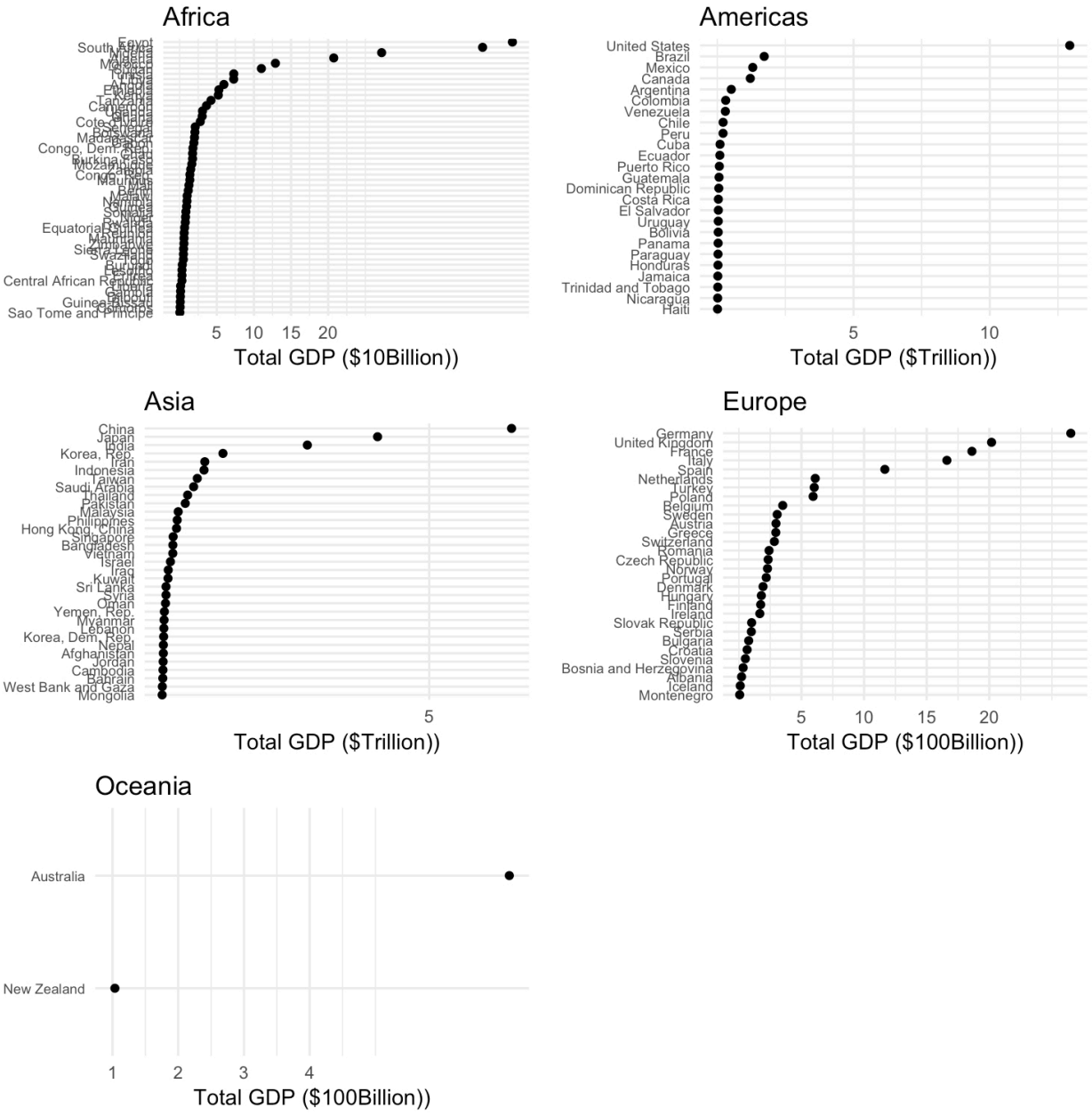


Figure 3. Graphs from position–angle experiment.

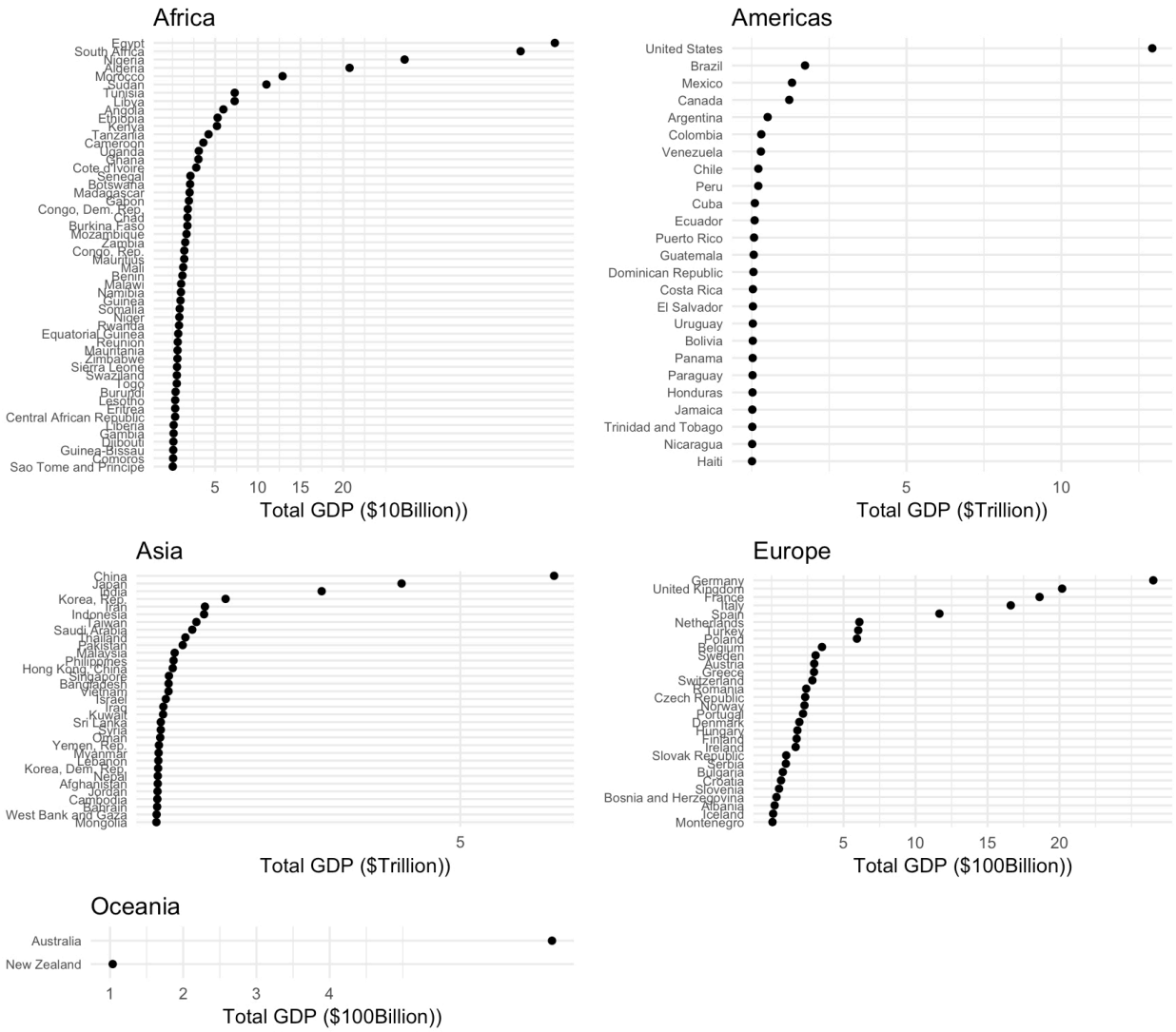
2. Gapminder Dashboard

a. create dashboard



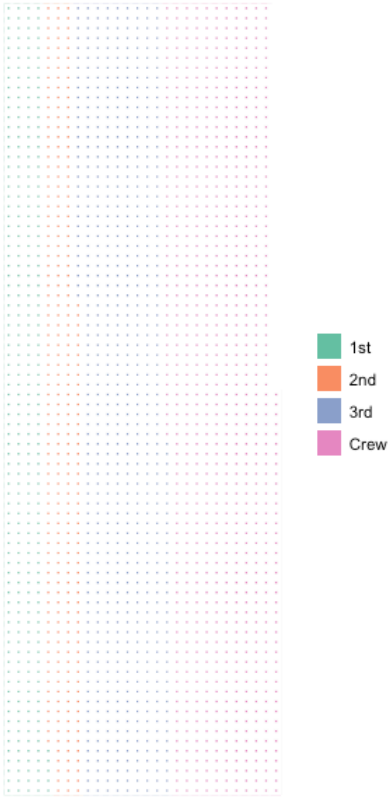
Data source: Jennifer Bryan (2015) Gapminder package

b. configure dashboard



Data source: Jennifer Bryan (2015) Gapminder package

3. Titanic Waffle Chart



Titanic Passengers by Class

CODE:

```
# Author: Brian Lambert
# Name: module_5_homework_code.R
# Description: Graphics exploring the cut of diamonds in the "diamonds" data
set,
# as well as comparing the cut to diamond size (carats).

# setwd("/Users/brianlambert/Desktop/STA404/Module_5")
library(tidyverse)
library(scales)


#========================== Data ==========================

library(gapminder)

myGapData <- gapminder %>%
  mutate(TotalGDP = pop*gdpPercap) %>%
  mutate(order_continent = factor(continent,
                        levels=c("Oceania", "Africa", "Europe", "Americas", "Asia")))
```

```
GDPsummaryDF <- myGapData %>%
  group_by(continent, year) %>%
  summarise(ContinentTotalGDP = sum(TotalGDP), ncountries = n())
GDPyearDF <- myGapData %>%
  group_by(year) %>%
  summarise(YearTotalGDP = sum(TotalGDP))
GDPcombo <- left_join(GDPsummaryDF, GDPyearDF, by="year")
GDPcombo <- GDPcombo %>% mutate(PropWorldGDP = ContinentTotalGDP /
YearTotalGDP,
                      PctWorldGDP = 100*PropWorldGDP,
                      ContGDPBillions = ContinentTotalGDP/1000000000)
GDPcombo <- GDPcombo %>%
  mutate(order_continent = factor(continent,
                        levels=c("Oceania", "Africa", "Europe", "Americas", "Asia")))

#========================== Graphic 1 ==========================

# Africa dot plot
africa_dot <- myGapData %>%
  filter(continent=="Africa") %>%
  filter(year==2007) %>%
  ggplot(aes(TotalGDP, fct_reorder(country,TotalGDP))) +
  geom_point() +
  theme(text = element_text(6),
      axis.text.y = element_text(6)) +
  scale_y_discrete(name=element_blank()) +
  scale_x_continuous(name="Total GDP ($10Billion))",
                breaks=10000000000*c(5,10,15,20),
                labels = c("5","10","15","20"))+
  labs(title="Africa") +
  theme_minimal() +
  theme(axis.text.y=element_text(size=7)) +
  theme(legend.position = "none")


# Americas dot plot
americas_dot <- myGapData %>%
  filter(continent=="Americas") %>%
  filter(year==2007) %>%
  ggplot(aes(TotalGDP, fct_reorder(country,TotalGDP))) +
  geom_point() +
  theme(text = element_text(6),
      axis.text.y = element_text(6)) +
```

```
  scale_y_discrete(name=element_blank()) +
  scale_x_continuous(name="Total GDP ($Trillion))",
                     breaks=1000000000000*c(5,10,15,20),
                     labels = c("5","10","15","20"))+
  labs(title="Americas") +
  theme_minimal() +
  theme(axis.text.y=element_text(size=7)) +
  theme(legend.position = "none")


# Asia dot plot
asia_dot <- myGapData %>%
  filter(continent=="Asia") %>%
  filter(year==2007) %>%
  ggplot(aes(TotalGDP, fct_reorder(country,TotalGDP))) +
  geom_point() +
  theme(text = element_text(6),
      axis.text.y = element_text(6)) +
  scale_y_discrete(name=element_blank()) +
  scale_x_continuous(name="Total GDP ($Trillion))",
                     breaks=1000000000000*c(5,10,15,20),
                     labels = c("5","10","15","20"))+
  labs(title="Asia") +
  theme_minimal() +
  theme(axis.text.y=element_text(size=7)) +
  theme(legend.position = "none")


# Europe dot plot
europe_dot <- myGapData %>%
  filter(continent=="Europe") %>%
  filter(year==2007) %>%
  ggplot(aes(TotalGDP, fct_reorder(country,TotalGDP))) +
  geom_point() +
  theme(text = element_text(6),
      axis.text.y = element_text(6)) +
  scale_y_discrete(name=element_blank()) +
  scale_x_continuous(name="Total GDP ($100Billion))",
                     breaks=100000000000*c(5,10,15,20),
                     labels = c("5","10","15","20"))+
  labs(title="Europe") +
  theme_minimal() +
  theme(axis.text.y=element_text(size=7)) +
  theme(legend.position = "none")
```

```
# Oceania dot plot
oceania_dot <- myGapData %>%
  filter(continent=="Oceania") %>%
  filter(year==2007) %>%
  ggplot(aes(TotalGDP, fct_reorder(country,TotalGDP))) +
  geom_point() +
  theme(text = element_text(6),
      axis.text.y = element_text(6)) +
  scale_y_discrete(name=element_blank()) +
  scale_x_continuous(name="Total GDP ($100Billion))",
                breaks=100000000000*c(1,2,3,4),
                labels = c("1","2","3","4"))+
  labs(title="Oceania", caption="Data source: Jennifer Bryan (2015) Gapminder
package") +
  theme_minimal() +
  theme(axis.text.y=element_text(size=7)) +
  theme(legend.position = "none")

# dashboard of gdp for each country, grouped by continent in 2007
grid.arrange(africa_dot, americas_dot, asia_dot, europe_dot, oceania_dot)


#========================== Graphic 2 ==========================

# dashboard of gdp for each country, grouped by continent in 2007, height
adjusted for # countries
grid.arrange(africa_dot, americas_dot, asia_dot, europe_dot, oceania_dot,
          ncol = 2, nrow =3, heights = c(3,2,1))


#========================== Graphic 3 ==========================

library(waffle)
library(ggmosaic)
library(dplyr)

library(grid)
pushViewport(viewport(layout=grid.layout(2,2)))

myTitanic <- as.data.frame(Titanic)
titanic_class <- myTitanic %>%
            group_by(Class) %>%
```

```
            summarize(passengers = sum(Freq))

count_passengers <- titanic_class$passengers
names(count_passengers) <- as.vector(titanic_class$Class)
waffle(parts=count_passengers,
    rows = 80,
    xlab="Titanic Passengers by Class")
```