r_solution.R

```r
library(ggplot2)
library(dplyr)



mobiles_dataset <- read.csv("mobiles_dataset.csv")
mobiles_dataset <- mobiles_dataset %>%
  mutate(Launched.Price.Pakistan.PKR = Launched.Price.Pakistan.PKR*0.0036)
mobiles_dataset <- mobiles_dataset %>%
  mutate(Launched.Price.India.INR = Launched.Price.India.INR*0.011)
mobiles_dataset <- mobiles_dataset %>%
  mutate(Launched.Price.China.CNY = Launched.Price.China.CNY*0.14)
mobiles_dataset <- mobiles_dataset %>%
  mutate(Launched.Price.Dubai.AED = Launched.Price.Dubai.AED*0.27)

# PART 1:

# --1
selected_cols <- c("Battery.Capacity.mAh",
                   "Launched.Price.Pakistan.PKR",
                   "Launched.Price.India.INR",
                   "Launched.Price.China.CNY",
                   "Launched.Price.USA.USD",
                   "Launched.Price.Dubai.AED")

brands <- unique(mobiles_dataset$Company.Name)

for (brand in brands) {

  brand_data <- mobiles_dataset %>%
    filter(Company.Name == brand) %>%
    select(all_of(selected_cols))


  if (nrow(brand_data) > 1) {
    correlation_matrix <- cor(brand_data)
    print(brand)
    for (region in selected_cols[-1]) {
      cat(region, ":", correlation_matrix["Battery.Capacity.mAh", region], "\n")
    }
  }
}

# results are printed in the R console
# here You can see each brands correlation with currency and battery size, mostly increase of battery size will not increase the price
# google is an exception it has a kind of good correlation which means with battery size increase the price increases
# the error empty/NA values are for the brands that have only 1 Battery size option


# --2
mobiles_dataset <- mobiles_dataset %>%
  mutate(RAM = gsub("GB|/", "", RAM),
         RAM = gsub("8  12", "8", RAM),
         RAM = as.integer(RAM))

selected_cols <- c("RAM",
                   "Launched.Price.Pakistan.PKR",
                   "Launched.Price.India.INR",
                   "Launched.Price.China.CNY",
                   "Launched.Price.USA.USD",
                   "Launched.Price.Dubai.AED")


for (brand in brands) {

  brand_data <- mobiles_dataset %>%
    filter(Company.Name == brand) %>%
    select(all_of(selected_cols))


  if (nrow(brand_data) > 1) {
    correlation_matrix <- cor(brand_data)
    print(brand)
    for (region in selected_cols[-1]) {
      cat(region, ":", correlation_matrix["RAM", region], "\n")
    }
  }
}
# results are printed in the R console
# here You can see each brands correlation with currency and RAM, and mostly yes the with RAM increase the prices increases
# the error empty/NA plots are for the brands that has only 1 RAM option


# --3
price_columns <- c("Launched.Price.Pakistan.PKR",
                "Launched.Price.India.INR",
                "Launched.Price.China.CNY",
                "Launched.Price.USA.USD",
                "Launched.Price.Dubai.AED")

mobiles_dataset <- mobiles_dataset %>%
  rowwise() %>%
  mutate(price_variance = sd(c_across(all_of(price_columns)), na.rm = TRUE)) %>%
```

```r
  ungroup()


mean_variance <- mobiles_dataset %>%
  group_by(Company.Name) %>%
    summarise(price_variance_mean = mean(price_variance)) %>%
      arrange(price_variance_mean)
# as we can see the variance of prices in each region for each company averaged for their different mobile models
# The most variance has Huawei averaged for different device models, and the least is Nokia which is the most stable across different regions


# --4
mobiles_dataset <- mobiles_dataset %>%
  mutate(Category = case_when(
    Launched.Price.USA.USD <= 300 ~ "budget",
    Launched.Price.USA.USD > 300 & Launched.Price.USA.USD <= 700 ~ "mid-range",
    TRUE ~ "premium"
  ))

grouped_category_count <- mobiles_dataset %>%
  group_by(Company.Name, Category) %>%
    summarise(Count = n())
# here are the names of the brands and what category they produced


# --5
pakistan_mean <- mean(mobiles_dataset$Launched.Price.Pakistan.PKR)
india_mean <- mean(mobiles_dataset$Launched.Price.India.INR)
china_mean <- mean(mobiles_dataset$Launched.Price.China.CNY)
usa_mean <- mean(mobiles_dataset$Launched.Price.USA.USD)
dubai_mean <- mean(mobiles_dataset$Launched.Price.Dubai.AED)
print(which.min(c(pakistan_mean,india_mean,china_mean,usa_mean,dubai_mean))-1) # index of the lowest mean
# results are printed in the R console


# PART 2:

# --1
price_data <- data.frame(
  Region = c("Pakistan", "India", "China", "USA", "Dubai"),
  Average_Price = c(pakistan_mean, india_mean, china_mean, usa_mean, dubai_mean)
)
ggplot(data =price_data, aes(x=Average_Price, y=Region)) +
  geom_bar(stat = "identity")



# --2
market_share <- mobiles_dataset %>%
  group_by(Company.Name) %>%
  summarise(number_of_phones = n())

ggplot(market_share, aes(x = "", y = number_of_phones, fill = Company.Name)) +
geom_bar(stat = "identity", width = 1) +
coord_polar("y", start = 0) +
theme_void()




# PART 3:

# --1

ggplot(mobiles_dataset, aes(y=Launched.Price.USA.USD, x=Company.Name, fill = Company.Name)) +
  geom_boxplot() +
  geom_point(,alpha =0.5) +
  labs(title = "Price Distribution by Company in USA", subtitle = "A boxplot showing how the price varies by company, with individial data po
  theme_minimal()+
  theme(axis.text.x = element_text(angle = 45, hjust = 1))


# --2
ggplot(mobiles_dataset, aes(x=Battery.Capacity.mAh, y=Launched.Price.USA.USD, color=Company.Name, size = Battery.Capacity.mAh))+
  geom_point()+
  labs(title = "Battery Capacity vs. Price in USA", subtitle = "The relationship between battery capacity, price, and screen size across diff
   guides(size = "none") +
  theme_minimal()



# --2
top_mobiles_dataset <- mobiles_dataset %>%
  filter(Company.Name %in% c("Apple","Honor", "Oppo", "Samsung", "Vivo"))



ggplot(top_mobiles_dataset, aes(x=Battery.Capacity.mAh, y=Launched.Price.USA.USD, color=Screen.Size.inches, size = Battery.Capacity.mAh, shap
  geom_point(alpha=0.7)+
  scale_shape_manual(values=c("Apple" = 16,"Honor" = 17,"Oppo" = 18,"Samsung" = 15,"Vivo" = 16)) +
  labs(title = "Battery Capacity vs. Price for Top 5 Brands", subtitle = "Different Shape for each Brand, Color by Screen Size (USA)", x="Bat
  guides(size = "none", color="none") +
  theme_minimal()
```