HW3

Vanesa Avoyan

2025-03-22

Neccessary libraries

```r
library(ggplot2)
library(dplyr)
library(tidyr)
library(stringr)
```

Loading the dataset

```r
mobiles_dataset <- read.csv("mobiles_dataset.csv")
```

Part 1.

```r
mobiles_dataset$Launched.Price.Pakistan.PKR <-
  mobiles_dataset$Launched.Price.Pakistan.PKR * 0.0036
mobiles_dataset$Launched.Price.India.INR <-
  mobiles_dataset$Launched.Price.India.INR * 0.011
mobiles_dataset$Launched.Price.China.CNY <-
  mobiles_dataset$Launched.Price.China.CNY * 0.14
mobiles_dataset$Launched.Price.Dubai.AED <-
  mobiles_dataset$Launched.Price.Dubai.AED * 0.27
```

```r
Mobile1 <- mobiles_dataset %>%
  pivot_longer(cols = starts_with("Launched.Price."),
               names_to = "Country",
               values_to = "Price") %>%
  mutate(Country = gsub("Launched.Price_", "", Country)) %>%
  mutate(Country = str_extract(Country, "(China|Dubai|India|Pakistan|USA)\\.[A-Z]{3}"))
```

  1)

```r
Mobile1.1 <- Mobile1 %>%
  group_by(Country, Battery.Capacity.mAh) %>%
  summarise(Std_of_Price = sd(Price, na.rm = TRUE), .groups = "drop") %>%
  filter(!is.na(Std_of_Price))
```

```r
ggplot(data = Mobile1.1, aes(x = Battery.Capacity.mAh,
                              y = Std_of_Price,
                              colour = Country)) +
  geom_point() +
  facet_grid(~Country) +
  labs(title = "Battery Capacity VS Price",
       x = "Battery Capacity",
       y = "Price") +
  theme_minimal()
```
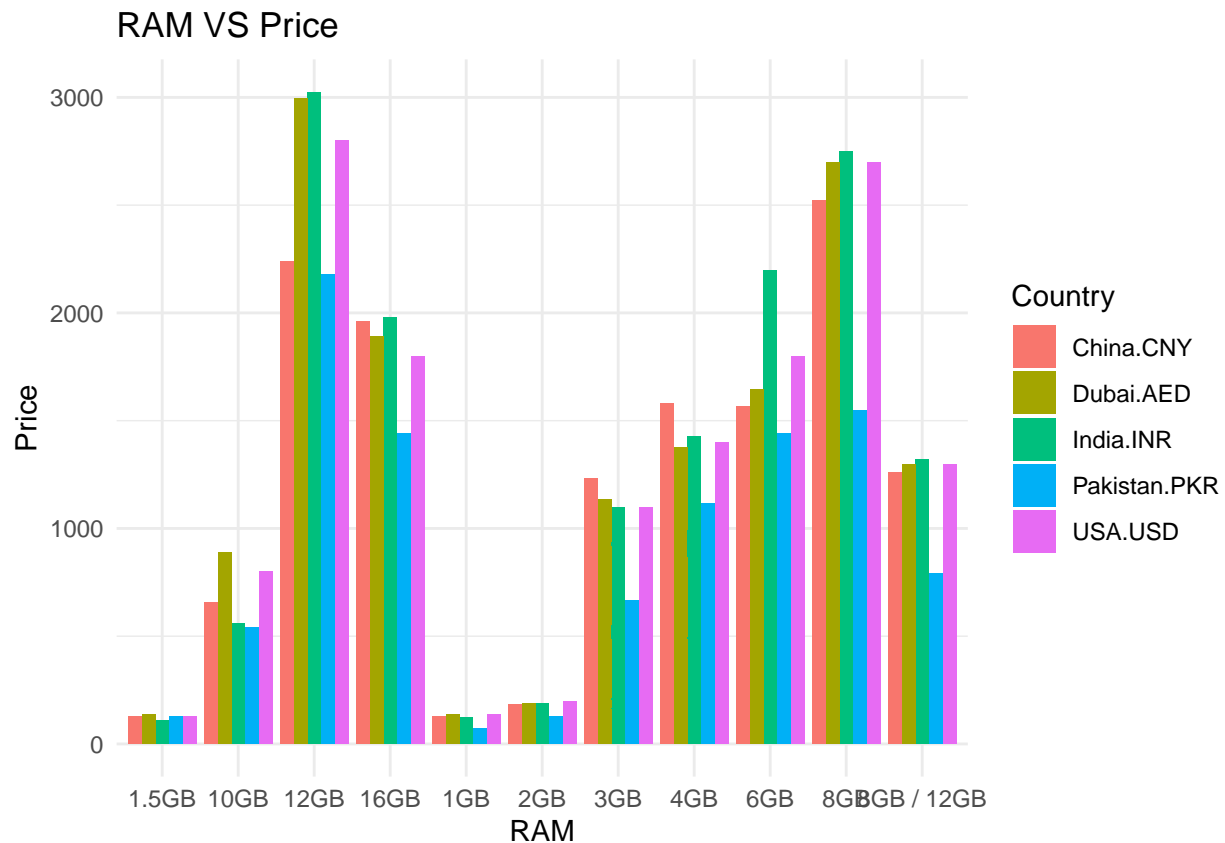
Battery Capacity VS Price
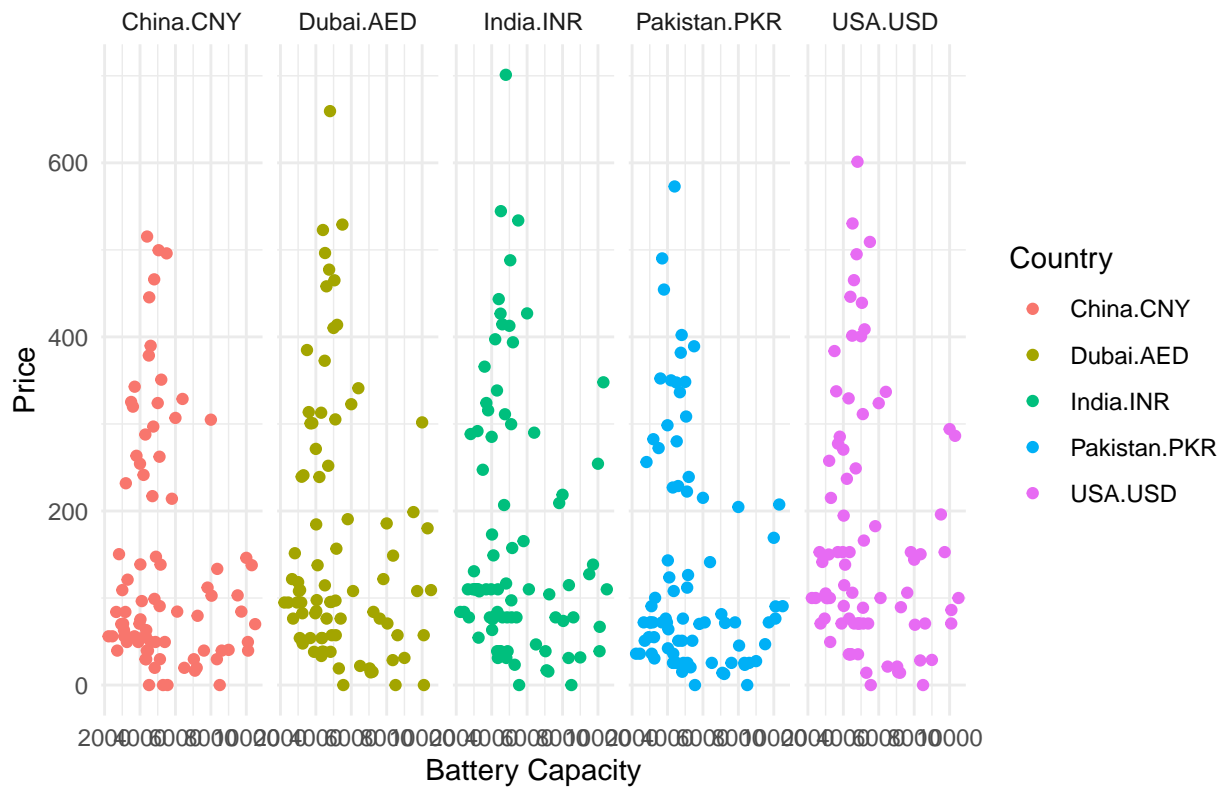
The Capacity influence is nearly the same.

2)

```
ggplot(data = Mobile1, aes(x = RAM,
                           y = Price,
                           fill = Country)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "RAM VS Price",
       x = "RAM",
       y = "Price") +
  theme_minimal()
```
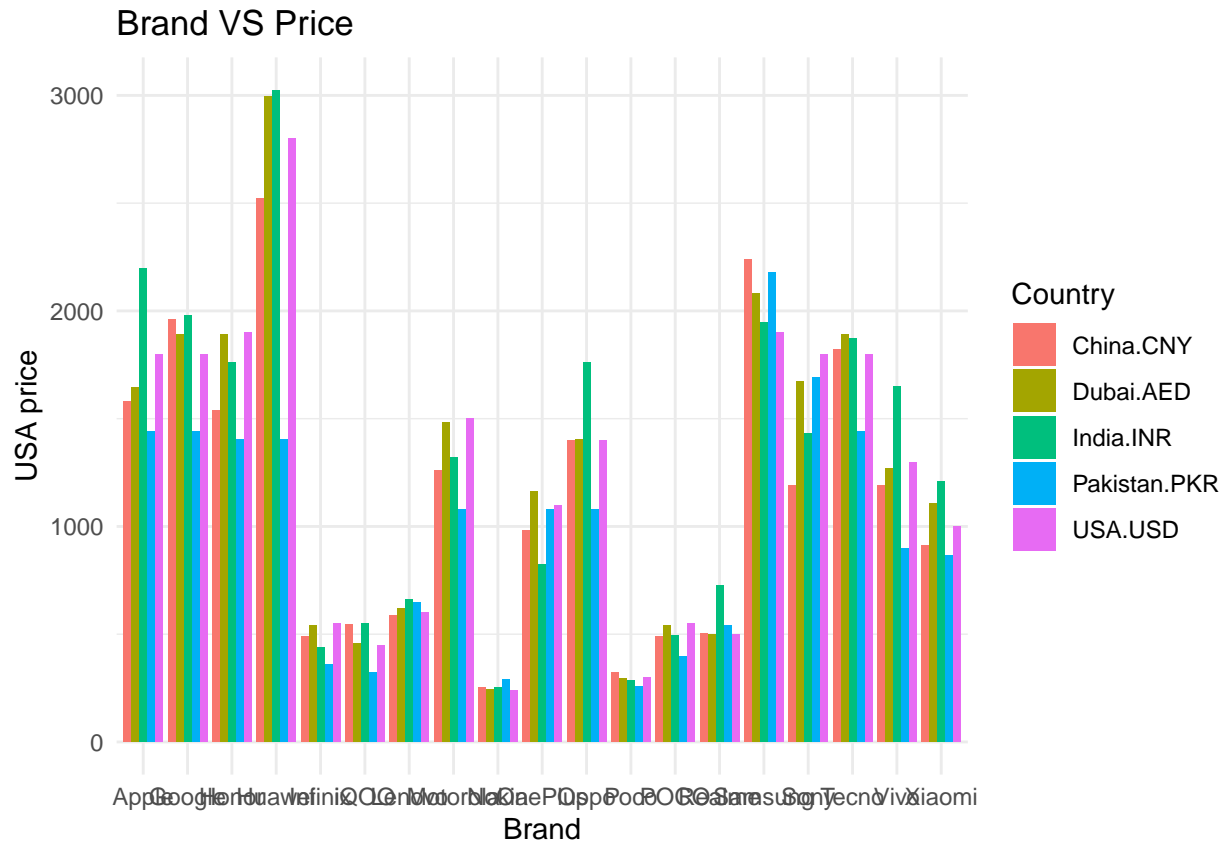
RAM VS Price

Yes. The RAM with 12 GB and 8 GB are the most expensive ones. Through the Pakistan prices are significantly lower.

3)

```
ggplot(data = Mobile1, aes(x = Company.Name,
                           y = Price,
                           fill = Country)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Brand VS Price",
       x = "Brand",
       y = "USA price") +
  theme_minimal()
```

Brand VS Price

Not only Apple has big variance but also other brands. For the Apple India has the highest markup. There are Poco, Nokia, Lenovo that have more stable pricing policy.

4)

```r
Mobile1$Affordability <- c("1")
```

```r
for (i in 1:nrow(Mobile1)) {
  if (Mobile1$Price[i] > 700) {
    Mobile1$Affordability[i] = "Premium"
  } else if (Mobile1$Price[i] < 300) {
    Mobile1$Affordability[i] = "Budget"
  } else (Mobile1$Affordability[i] = "Mid-range")
}
```

```r
ggplot(data = Mobile1, aes(x = Company.Name,
                           fill = Affordability)) +
  geom_bar(position = "fill") +
  labs(title = "Affordability Segment Distribution",
       x = "Affordability Segment",
       y = "Proportion") +
  theme_minimal()
```

Affordability Segment Distribution
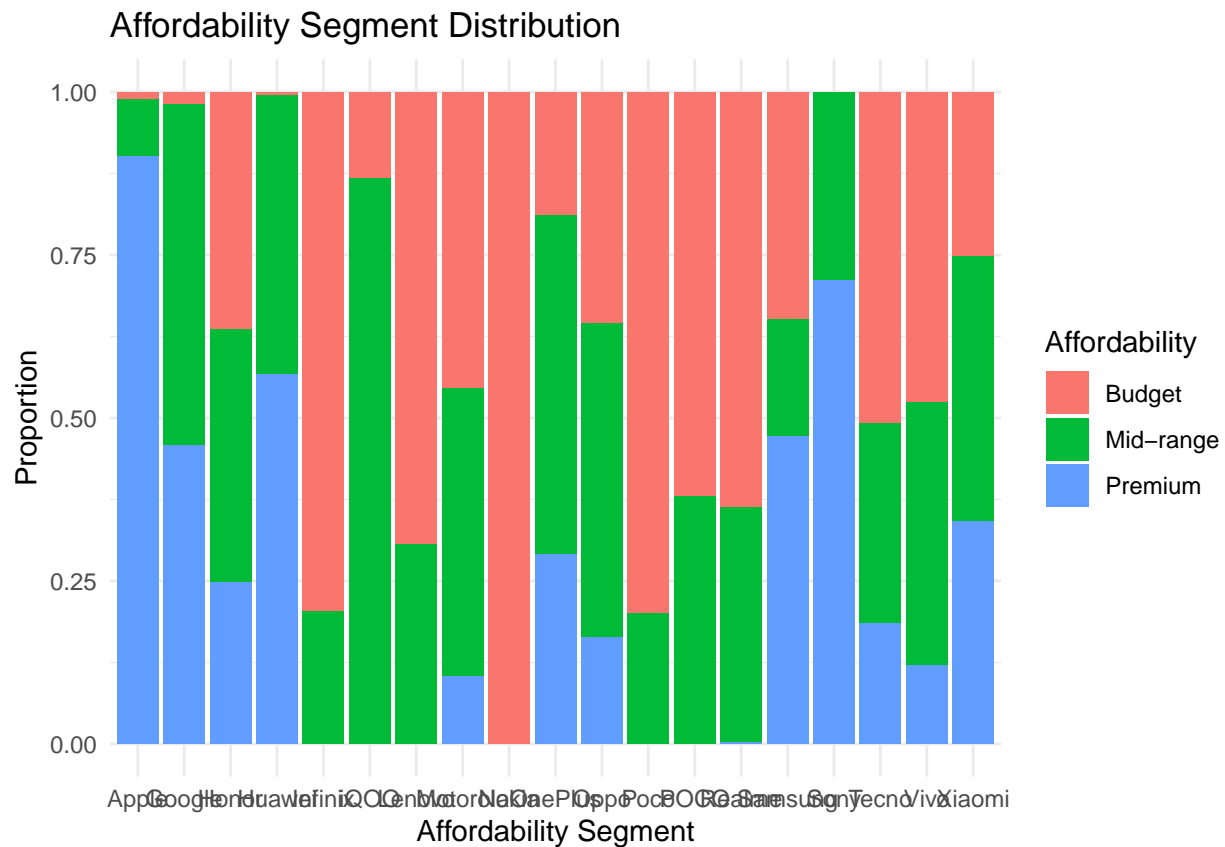

Here we can see that Nokia is only Budget priced phones. But brands like Apple, Sony and Huawei are Premium oriented.

5)

```
Mobile1.2 <- Mobile1 %>%
  group_by(Country, Affordability, Company.Name) %>%
  summarise(Avg_Price = mean(Price), .groups = "drop")
```

```
ggplot(data = Mobile1.2, aes(x = Company.Name,
                             fill = Affordability)) +
  geom_bar(position = "fill") +
  facet_wrap(~Country) +
  labs(title = "Affordability Segment Distribution by Country",
       x = "Affordability Segment",
       y = "Proportion") +
  theme_minimal()
```

Affordability Segment Distribution by Country
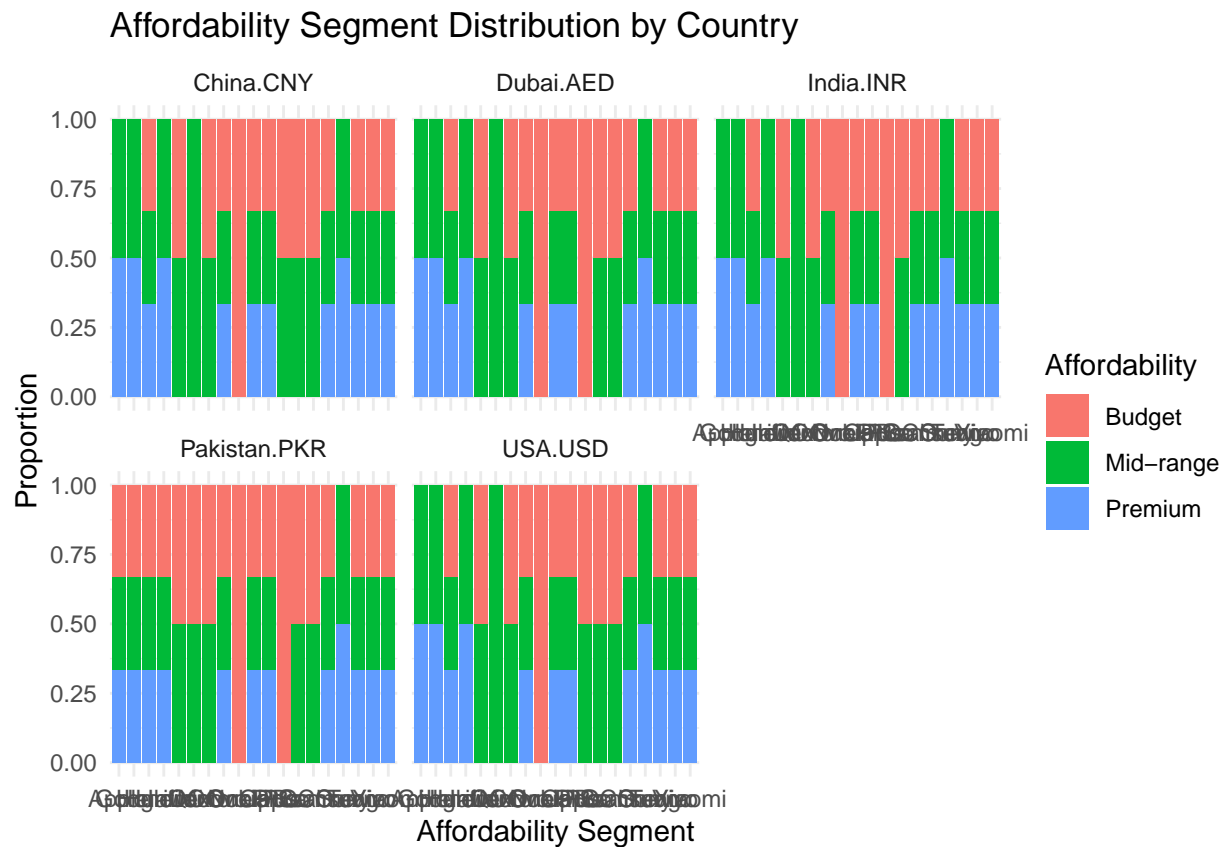
Pakistan offers the most affordable prices.

Part 2. 1)

```r
ggplot(data = Mobile1.2, aes(x = Country,
                             y = Avg_Price,
                             fill = Country)) +
  geom_bar(stat = "identity") +
  labs(title = "Bar chart of Countries average prices",
       x = "Country",
       y = "Price") +
  theme_minimal()
```
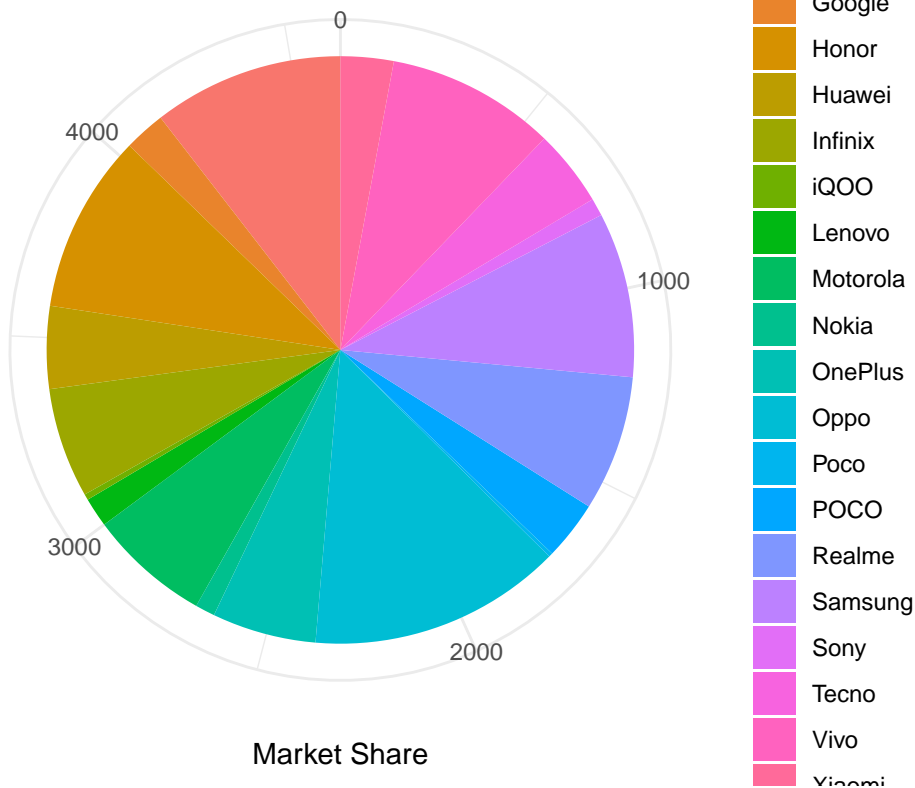
Bar chart of Countries average prices

2)
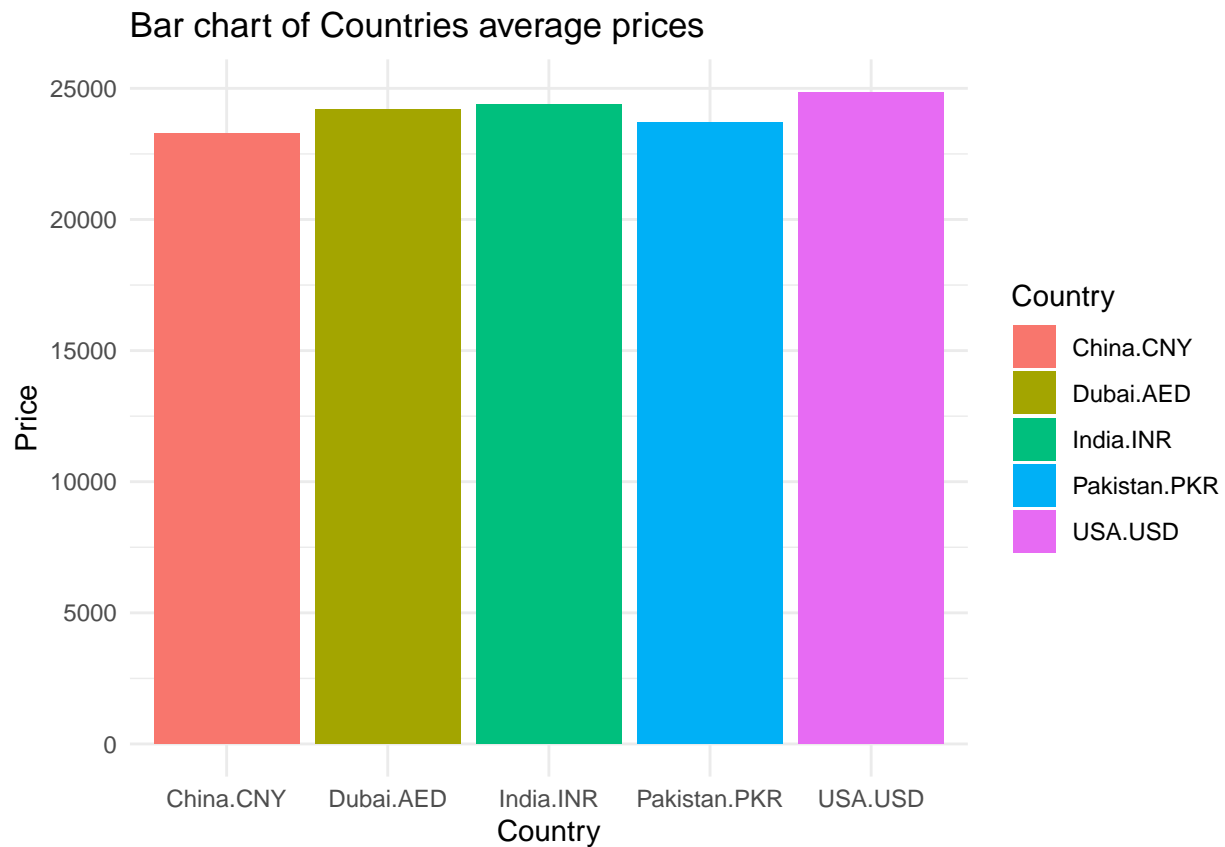
```r
Mobile2.1 <- Mobile1 %>%
  group_by(Company.Name) %>%
  summarise(Market_Share = n())
```

```r
ggplot(data = Mobile2.1, aes(x = "",
                             y = Market_Share,
                             fill = Company.Name)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  labs(title = "Market Share by Company",
       x = "",
       y = "Market Share") +
  theme_minimal()
```

Market Share by Company

Legend:
- Apple
- Google
- Honor
- Huawei
- Infinix
- iQOO
- Lenovo
- Motorola
- Nokia
- OnePlus
- Oppo
- Poco
- POCO
- Realme
- Samsung
- Sony
- Tecno
- Vivo
- Xiaomi

Market Share

Part 3. 1)

```r
Mobile3.1 <- Mobile1 %>%
  filter(Country == "USA.USD")
```

```r
ggplot(data = Mobile3.1, aes(x = Company.Name,
                             y = Price,
                             fill = Company.Name)) +
  geom_boxplot() +
  geom_point() +
  labs(title = "Price distribution by Company in USA",
       subtitle = "A boxplot showing how the price variesby company, with individual data points overla
       x = "Company",
       y = "Price in USD",
       fill = "Company Name") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45),
        plot.title = element_text(face = "bold"),
        plot.subtitle = element_text(face = "italic"))
```

Price distribution by Company in USA

A boxplot showing how the price variesby company, with individual data points overlaid

2)
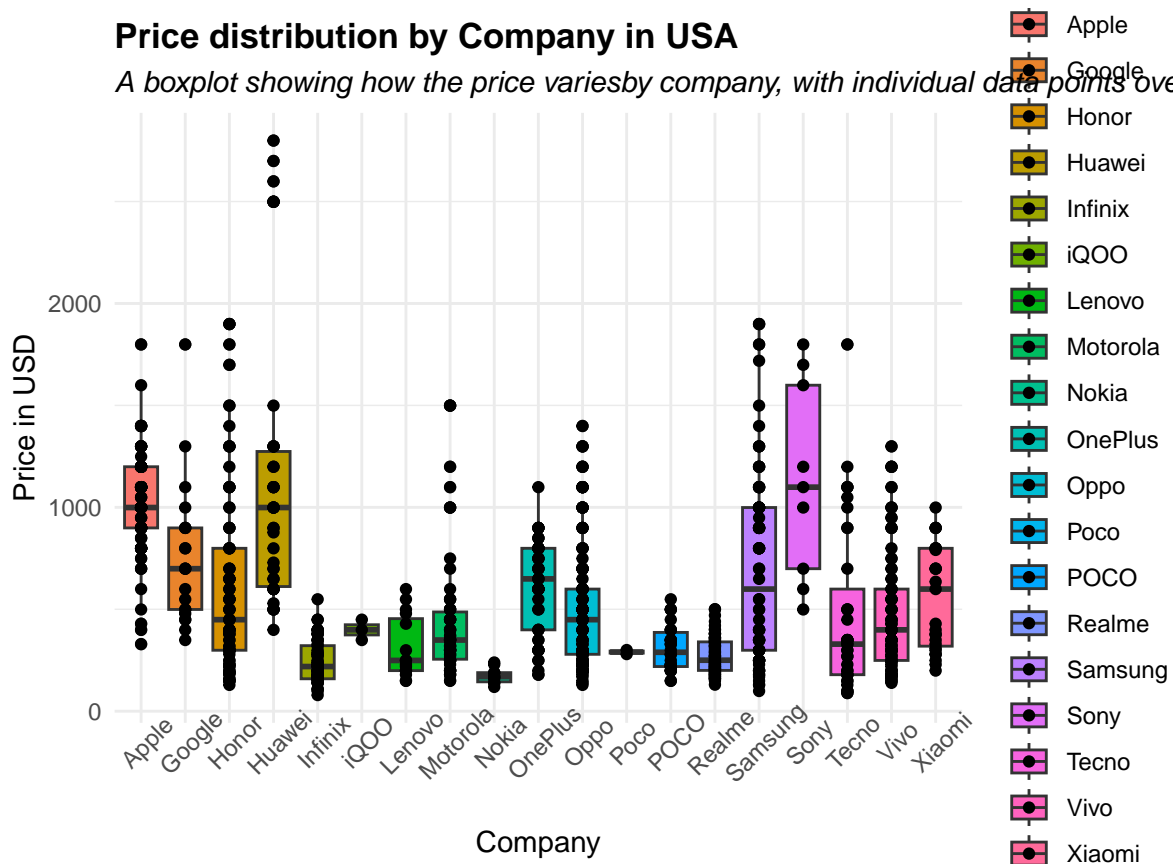
```r
ggplot(data = Mobile3.1, aes(x = Battery.Capacity.mAh,
                             y = Price,
                             colour = Company.Name,
                             size = Screen.Size.inches)) +
geom_point() +
labs(title = "Battery Capacity VS Price in USA",
     subtitle = "The relationship between capacity, price, and screen size across different smartphone
     x = "Battery Capacity",
     y = "Price",
     color = "Brand") +
theme_minimal() +
theme(
     plot.title = element_text(face = "bold"),
     plot.subtitle = element_text(face = "italic")) +
guides(size = "none")
```
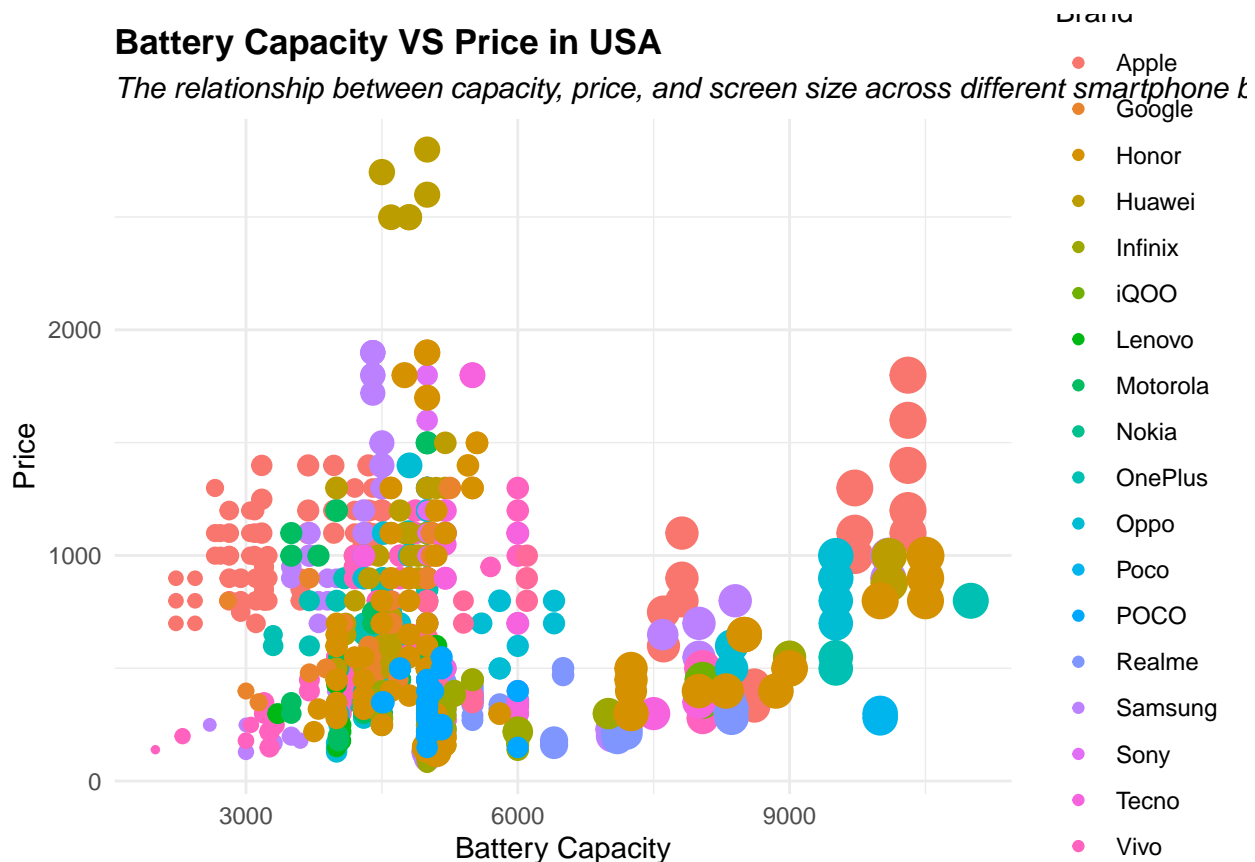
Battery Capacity VS Price in USA

The relationship between capacity, price, and screen size across different smartphone b

Brand
- Apple
- Google
- Honor
- Huawei
- Infinix
- iQOO
- Lenovo
- Motorola
- Nokia
- OnePlus
- Oppo
- Poco
- POCO
- Realme
- Samsung
- Sony
- Tecno
- Vivo

3)

```r
Top_Brands <- c("Apple", "Honor", "Oppo", "Samsung", "Vivo")
Mobile3.2 <- Mobile1 %>%
  filter(Company.Name %in% Top_Brands)
```

```r
ggplot(data = Mobile3.2, aes(x = Battery.Capacity.mAh,
                             y = Price,
                             colour = Screen.Size.inches,
                             size = Battery.Capacity.mAh,
                             shape = Company.Name)) +
  geom_point(alpha = 0.7) +
  labs(title = "Battery Capacity VS Price for Top 5 Brands",
       subtitle = "Different Shapes for Each Brand, Color by Screen Size, (USA)",
       x = "Battery Capacity",
       y = "Price",
       shape = "Brand") +
  theme_minimal() +
  theme(
        plot.title = element_text(face = "bold"),
        plot.subtitle = element_text(face = "italic")) +
  guides(colour = "none", size = "none") +
  scale_shape_manual(values = c(16, 17, 18, 15, 19))
```

Battery Capacity VS Price for Top 5 Brands

Different Shapes for Each Brand, Color by Screen Size, (USA)