Tesla Project

Sreng Tang

2025-08-12

```r
# libraries
library(dplyr)
library(tidyr)
library(ggplot2)
library(scales)
library(paletteer)

# read in the data
df <- read.csv("Data/Electric_Vehicle_Population_Data.csv")
```

```r
# data cleaning
clean_df <- df %>%
  rename_all(~ gsub("\\.","_", .)) %>% # gsub = global substitute
    mutate(
      Base_MSRP = as.numeric(Base_MSRP),
      Tesla = ifelse(Make == "TESLA", "TESLA", "Other")
    )
```

1. Tesla Market Share in Washington (%)

```r
total_vehicles <- nrow(clean_df)
tesla_vehicles <- clean_df %>%
  filter(Tesla == "TESLA") %>%
  nrow()

tesla_market_share <- round(tesla_vehicles/total_vehicles * 100, 1)

market_df <- data.frame(
  group = c("TESLA", "OTHER"),
  market_share = c(tesla_market_share, 100 - tesla_market_share)
)

market_df <- market_df %>%
  mutate(
    label = paste0(market_share, "%"),
    ypos = cumsum(market_share) - 0.5 * market_share
  )
```
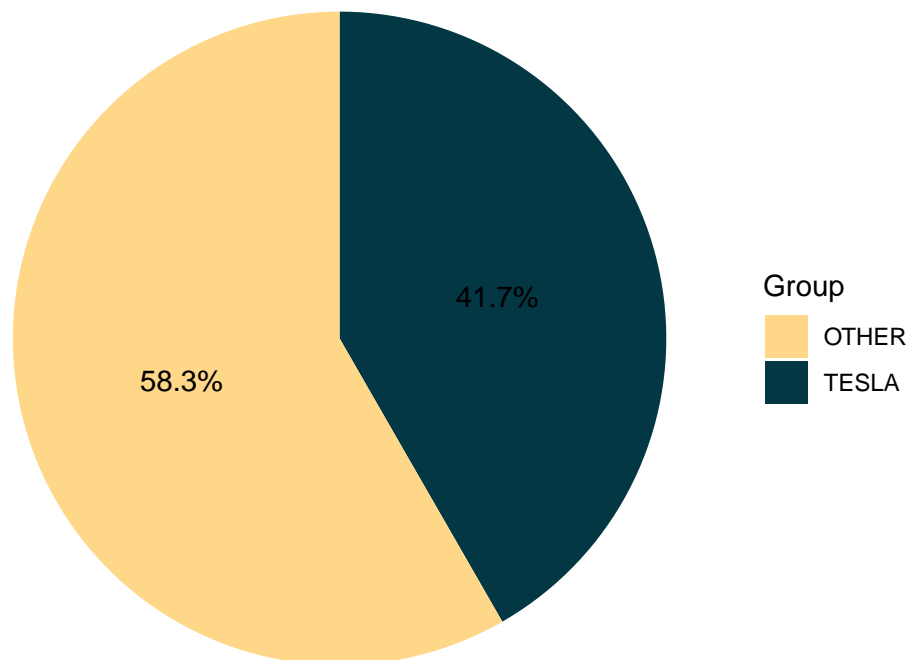
```
ggplot(market_df, aes(x = "", y = market_share, fill = group)) +
  geom_col() +
  coord_polar("y") +
  geom_text(aes(y = ypos, label = label)) +
  scale_fill_paletteer_d("nationalparkcolors::Acadia") +
  labs(title = "Tesla Market Share in Washington (%)", fill = "Group") +
  theme(legend.title = element_blank()) +
  theme_void()
```

Tesla Market Share in Washington (%)

2. Top Tesla Models Selling in the area - How much are we making from these models?

```
make_model_sold <- clean_df %>%
  filter(Tesla == "TESLA") %>%
  group_by(Make, Model) %>%
  summarise(count_sold = n(), .groups = "drop") %>%
  arrange(desc(count_sold))
```
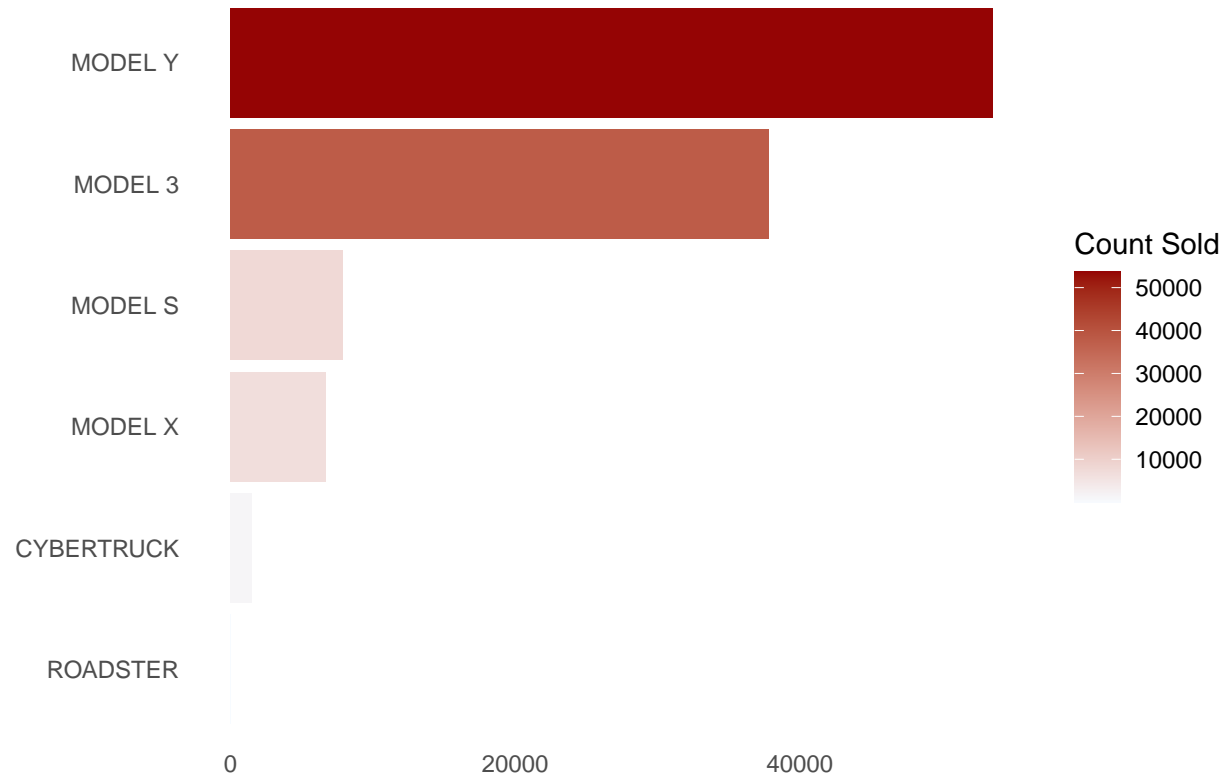
```
ggplot(make_model_sold, aes(x = reorder(Model, count_sold), y = count_sold, fill = count_sold)) +
  geom_col() +
  coord_flip() +
  labs(title = "Tesla Models Sold - All time", fill = "Count Sold",
       x = NULL,
```

```
      y = NULL) +
  theme_minimal() +
  theme(panel.grid = element_blank()) +
  scale_fill_gradient(
    low = "#F7FBFF",
    high = "#950404FF"
  )
```

Tesla Models Sold – All time


```
msrp_by_year <- clean_df %>%
  filter(Base_MSRP > 0) %>%
  filter(Tesla == "TESLA") %>%
  group_by(Model_Year, Make, Model) %>%
  summarise(min(Base_MSRP)) %>%
  arrange(Model_Year, Make, Model)
```

```
df_tesla_prices <- read.csv("Data/Tesla_Current_Base_Prices.csv")

# transform data to upper
df_tesla_prices <- df_tesla_prices %>%
  mutate(Model = toupper(Model))

# getting our numbers
top_tesla_model_price <- make_model_sold %>%
  left_join(df_tesla_prices, by = "Model") %>%
```

```r
  mutate(
    estimate_revenue = as.numeric(count_sold) * as.numeric(Base_Price_USD)
  )

# final estimate for lifetime sales
top_tesla_model_price <- top_tesla_model_price %>%
  filter(!is.na(estimate_revenue)) %>%
  group_by(Model) %>%
  slice_min(Base_Price_USD) %>%
  ungroup()
```

```r
# estimate revenue per model
ggplot(top_tesla_model_price, aes(x = reorder(Model, estimate_revenue), y = estimate_revenue, fill = est
  geom_col() +
  coord_flip() +
  labs(title = "Estimated Revenue Per Model", fill = "Estimated Revenue",
       x = NULL,
       y = NULL) +
  scale_y_continuous(labels = label_comma()) +
  scale_y_continuous(labels = label_number(scale = 1e-6, suffix = "M")) +
  theme_minimal() +
  theme(panel.grid = element_blank()) +
  scale_fill_gradient(
    low = "#F7FBFF",
    high = "#950404FF",
    labels = label_number(scale = 1e-6, suffix = "M")
  )
```

Estimated Revenue Per Model
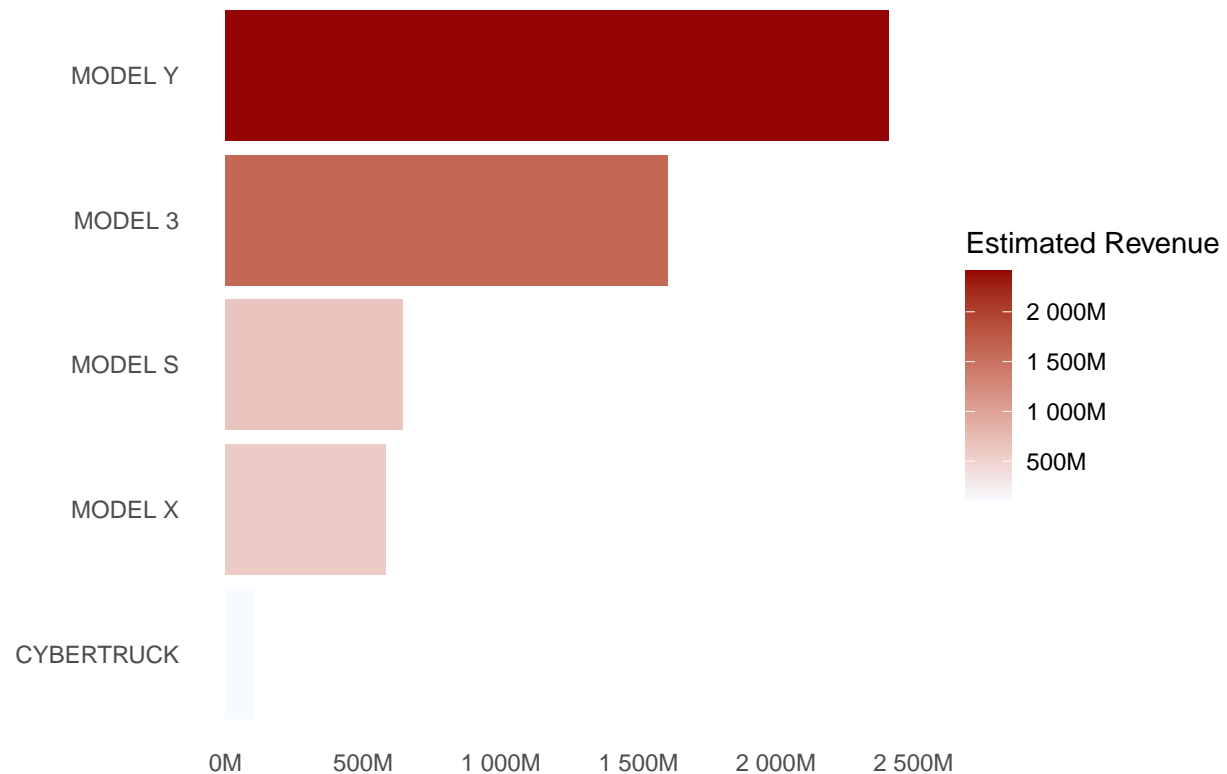


```r
# Tesla vehicle sold by year data
tesla_sold_by_year <- clean_df %>%
  filter(Tesla == "TESLA") %>%
  group_by(Model_Year) %>%
  summarise(Count = n()) %>%
  arrange(Model_Year)
```

```r
ggplot(tesla_sold_by_year, aes(x = Model_Year, y = Count)) +
  geom_line(color = "#023743FF", size = 1.1) +
  geom_point(color = "#E04B28FF", size = 2) +
  labs(title = "Tesla Vehicles Sold by Year",
       x = NULL,
       y = NULL)+
  theme_minimal() +
  theme(panel.grid = element_blank())
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

Tesla Vehicles Sold by Year
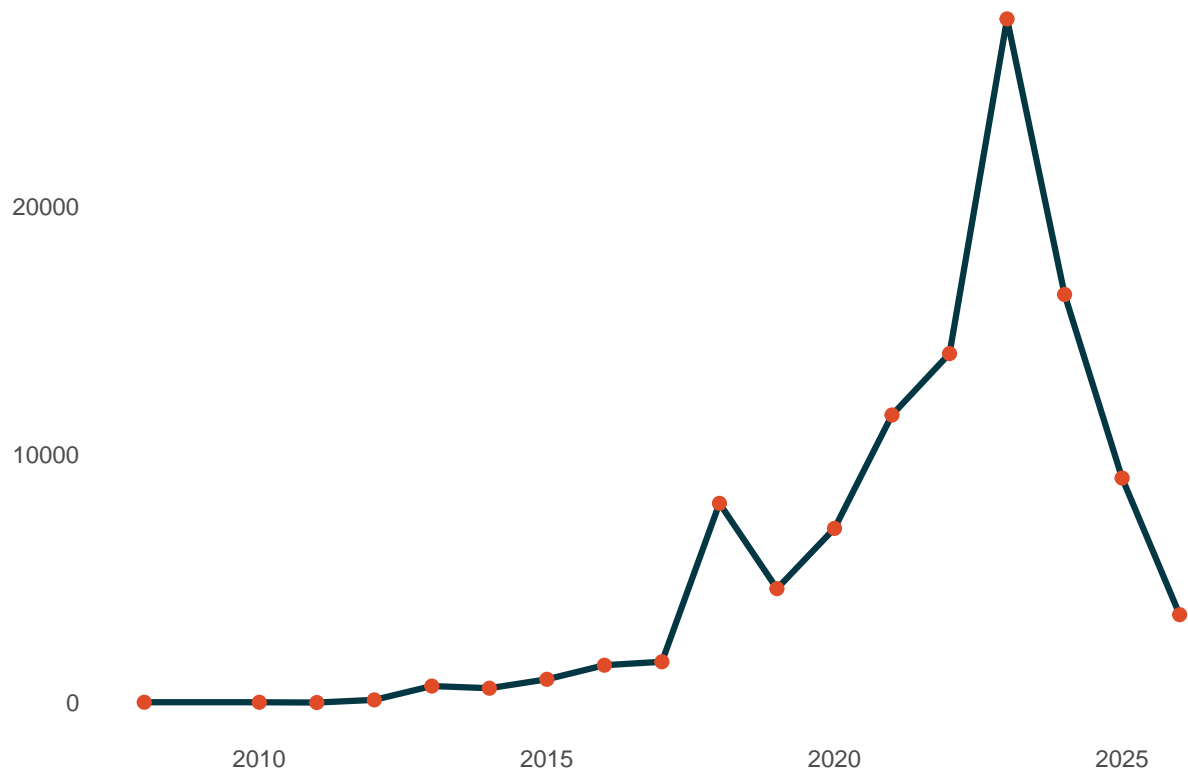


```r
# average range
mean_range_for_ev <- clean_df %>%
  filter(Electric_Vehicle_Type == "Battery Electric Vehicle (BEV)",
         Electric_Range != 0,
         Base_MSRP != 0) %>%
  group_by(Tesla) %>%
  summarise(avg_range = mean(Electric_Range),
            ave_msrp = mean(Base_MSRP))
```

```r
# median range
median_range_for_ev <- clean_df %>%
  filter(Electric_Vehicle_Type == "Battery Electric Vehicle (BEV)",
         Electric_Range != 0,
         Base_MSRP != 0) %>%
  group_by(Tesla) %>%
  summarise(median_range = median(Electric_Range),
            median_msrp = median(Base_MSRP))
```

```r
# Range
ggplot(mean_range_for_ev, aes(x = Tesla, y = avg_range, fill = Tesla)) +
  geom_col() +
  labs(title = "Average Electric Range (BEV)",
       fill = NULL,
       x = NULL,
       y = "Avg Range") +
```

```
scale_fill_paletteer_d("nationalparkcolors::Acadia") +
theme_minimal() +
theme(panel.grid = element_blank())
```

Average Electric Range (BEV)
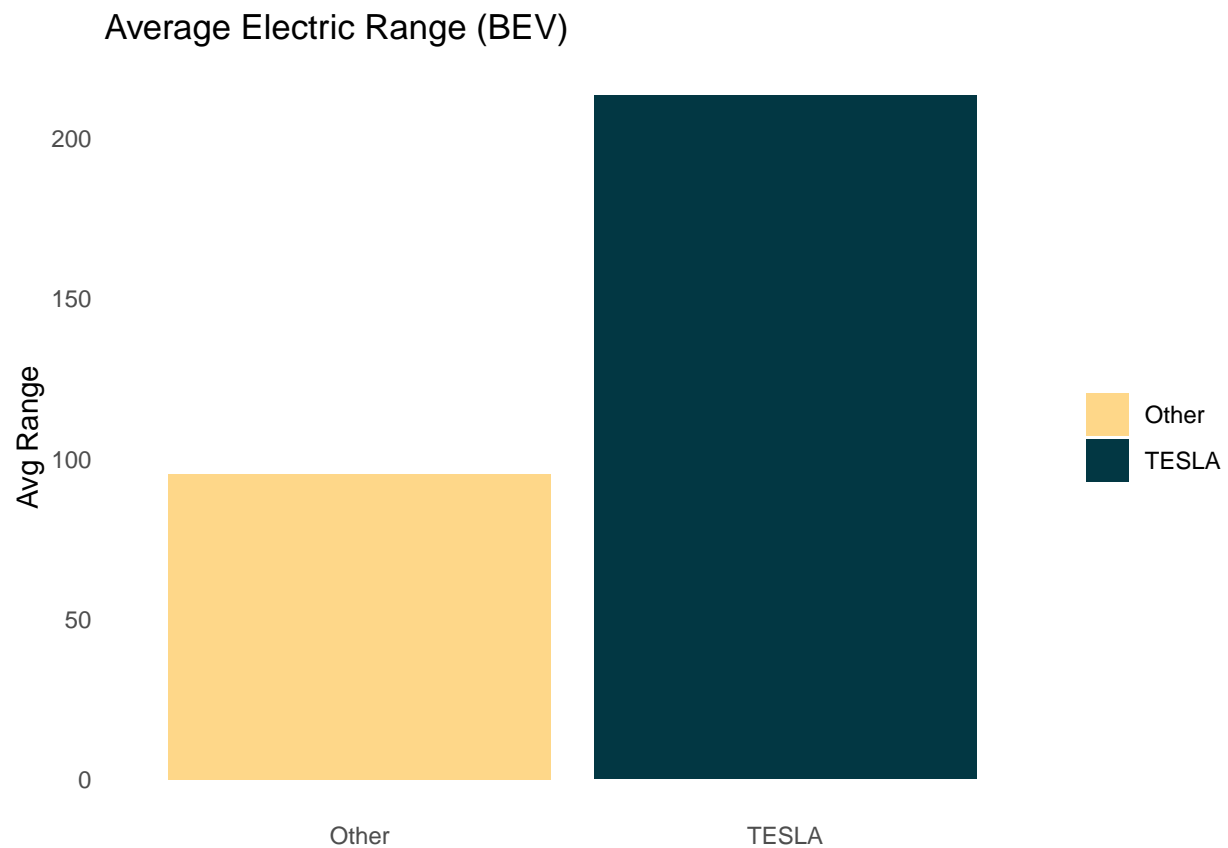

```
# MSRP
ggplot(mean_range_for_ev, aes(x = Tesla, y = ave_msrp, fill = Tesla)) +
  geom_col() +
  scale_fill_paletteer_d("nationalparkcolors::Acadia") +
  labs(title = "Average MSRP (BEV)",
       fill = NULL,
       x = NULL,
       y = "Avg MSRP") +
  theme_minimal() +
  theme(panel.grid = element_blank())
```

Average MSRP (BEV)
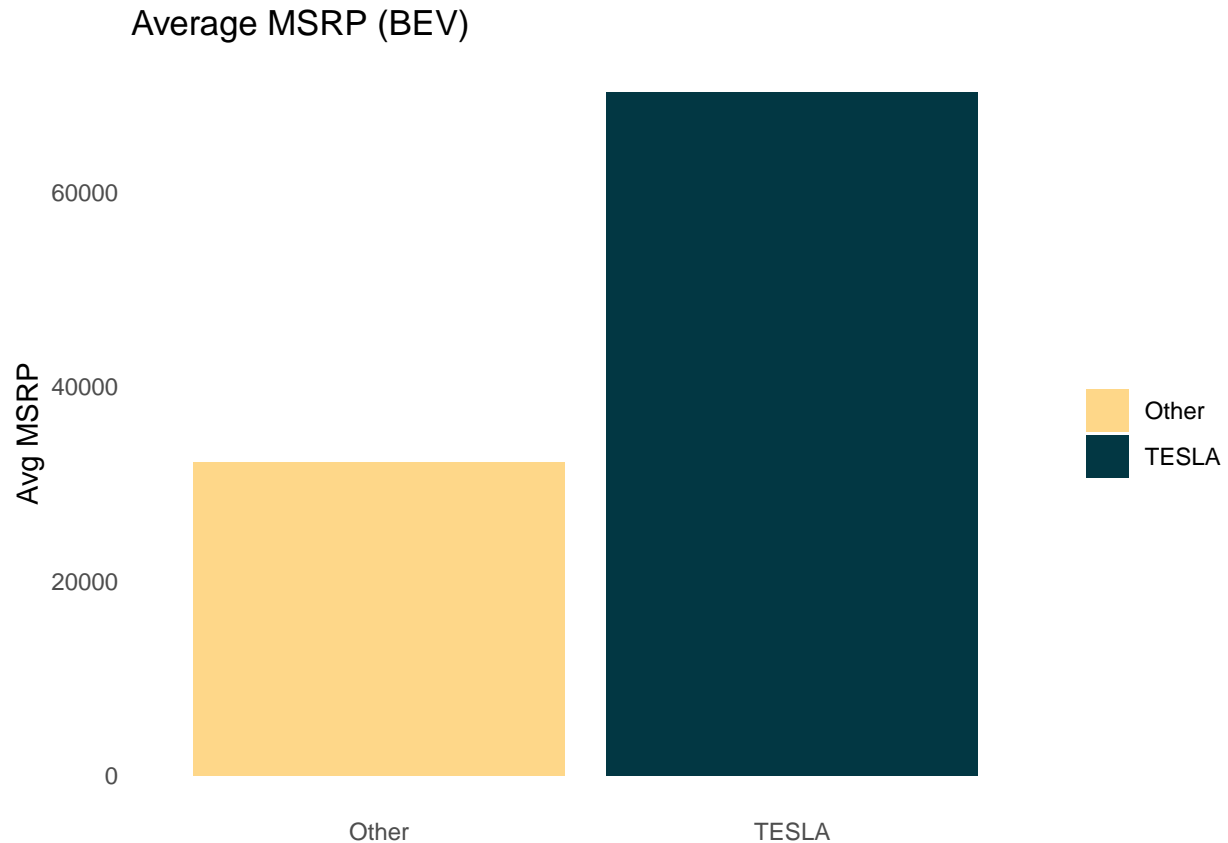
4. PHEV vs BEV Trends

```r
phev_vs_bev <- clean_df %>%
  group_by(Model_Year, Electric_Vehicle_Type) %>%
  summarise(Count = n(), .groups = "drop")
```

```r
ggplot(phev_vs_bev, aes(x = Model_Year, y = Count, color = Electric_Vehicle_Type)) +
  geom_line(size = 1.1) +
  geom_point(size = 2) +
  labs(title = "PHEV vs BEV Trends Over Time",
       color = NULL,
       x = NULL,
       y = NULL)  +
  theme_minimal() +
  scale_fill_manual() +
  theme(panel.grid = element_blank(),
        legend.position = "top",
        legend.justification = "left") +
  scale_color_manual(
  values = c(
      "Battery Electric Vehicle (BEV)" = "#950404FF",
      "Plug-in Hybrid Electric Vehicle (PHEV)" = "#0F542FFF"
    ))
```

PHEV vs BEV Trends Over Time

- Battery Electric Vehicle (BEV)
- Plug-in Hybrid Electric Vehicle (PHEV)
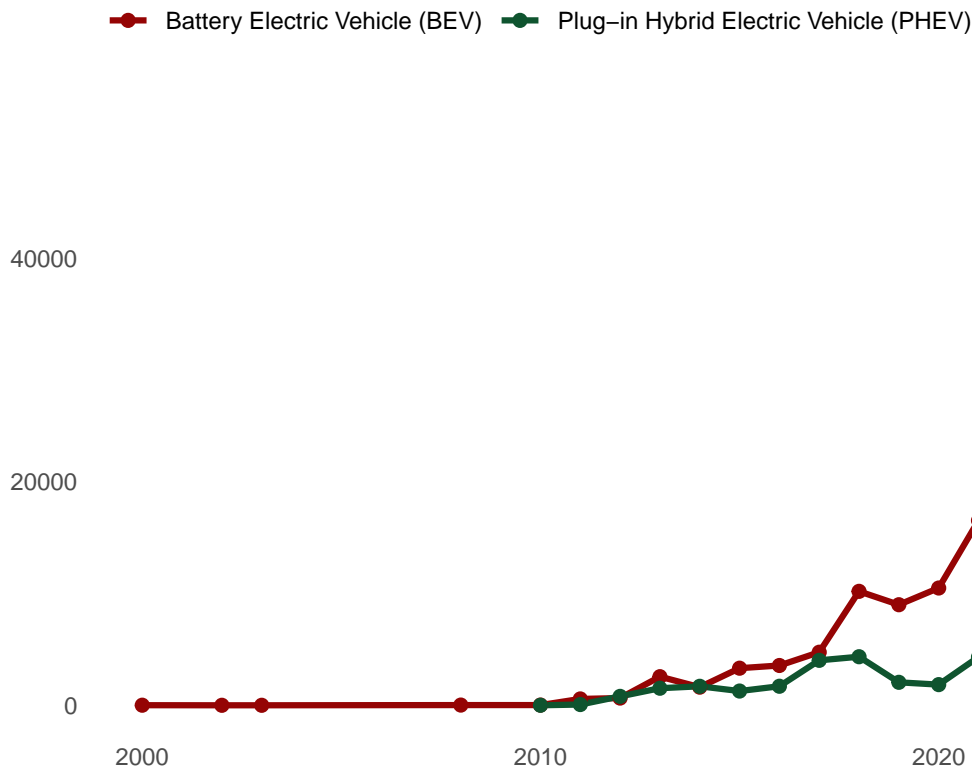
5. Top Electric Utilities in Washington for Tesla

```r
utility_count_for_tesla <- clean_df %>%
  filter(Tesla == "TESLA") %>%
  group_by(Electric_Utility) %>%
  summarise(Count = n(), .groups = "drop") %>%
  arrange(desc(Count)) %>%
  head(5)
```

```r
utility_count_for_tesla$Electric_Utility <- recode(utility_count_for_tesla$Electric_Utility,
  "PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)" = "Puget Sound & Tacoma",
  "PUGET SOUND ENERGY INC" = "Puget Sound Energy",
  "CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)" = "Seattle & Tacoma",
  "BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF CLARK COUNTY - (WA)" = "Bonneville & Clark",
  "BONNEVILLE POWER ADMINISTRATION||CITY OF TACOMA - (WA)||PENINSULA LIGHT COMPANY" = "Bonneville & Pen
)
```

```r
ggplot(utility_count_for_tesla, aes(x = reorder(Electric_Utility, Count), y = Count)) +
  geom_col(fill = "#004042FF") +
  coord_flip() +
  labs(
    title = "Top Electric Utilities for Tesla in Washington",
    x = NULL,
```

```
  y = NULL) +
theme_minimal() +
theme(panel.grid = element_blank())
```

Top Electric Utilities for Tesla in Washington