Mario_kart_analysis

We're importing the dataset from the Mario Kart

```
options(digits = 5)
```

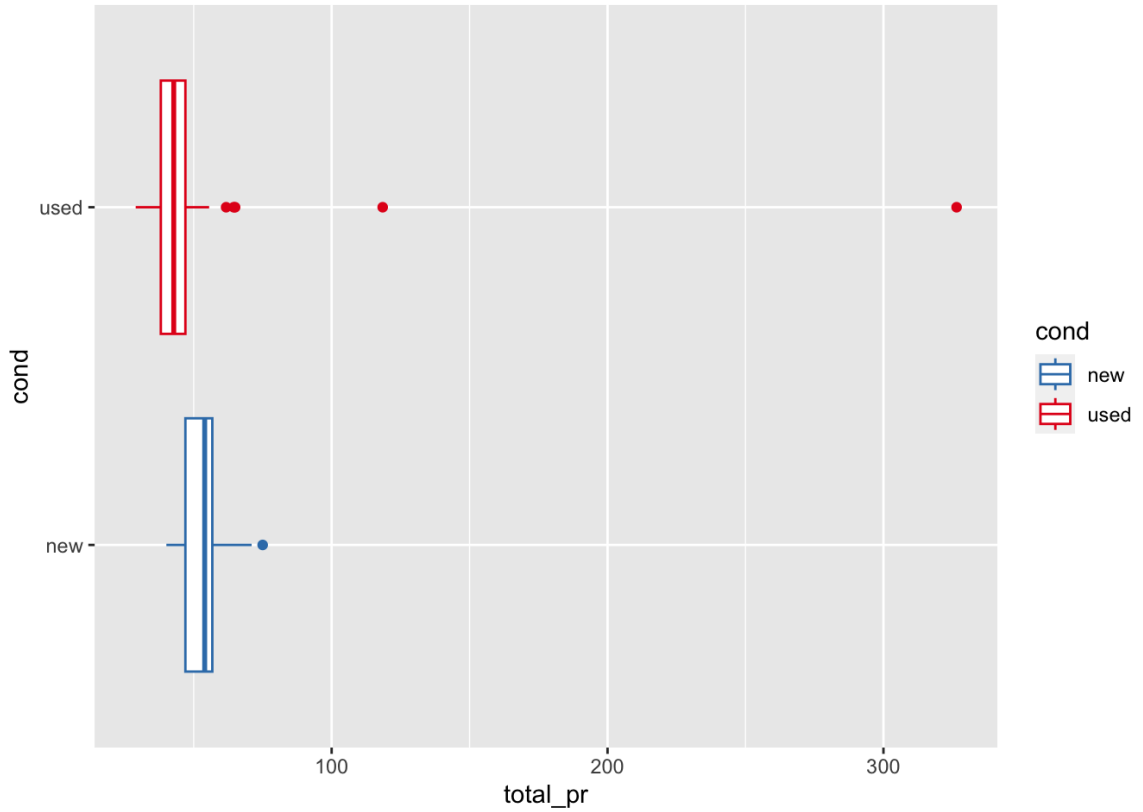
```
── Attaching core tidyverse packages ──────────────────────── tidyverse
2.0.0 ──
✔ dplyr     1.1.3     ✔ readr     2.1.4
✔ forcats   1.0.0     ✔ stringr   1.5.0
✔ ggplot2   3.4.4     ✔ tibble    3.2.1
✔ lubridate 1.9.2     ✔ tidyr     1.3.0
✔ purrr     1.0.2
── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()    masks stats::lag()
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
Loading required package: airports

Loading required package: cherryblossom

Loading required package: usdata
```

#This is our boxplot

```
mariokart |>
  ggplot(aes(x=cond, y=total_pr, color=cond)) +
  geom_boxplot() +
  coord_flip() +
  scale_color_brewer(palette = "Set1", direction = -1)
```




```
mario_kart_selected <-
  mariokart |>
  select(id, cond, total_pr)

mario_kart_selected <-
  arrange(mario_kart_selected,desc(total_pr))


used_mariokart <-
  mario_kart_selected |>
  filter(cond=="used")

used_mariokart<-
  arrange(used_mariokart,desc(total_pr))

used_mariokart <-
   filter(used_mariokart, !(total_pr %in% c(326.51, 118.5)))

used_mariokart<-
   arrange(used_mariokart,desc(total_pr))
```

```
new_mariokart <-
  mario_kart_selected |>
  filter(cond== "new")

new_mariokart <-
  arrange(new_mariokart,desc(total_pr))
```

Interquartile

#Here we removed the outliers. We removed the outliers which are 326.51, and 118.5

```
mario_kart_selected <-
    filter(mario_kart_selected, !(total_pr %in% c(326.51, 118.5)))
```

#This is our mean by group

```
mean_by_group <- mario_kart_selected |>
  group_by(cond) |>
  summarise(mean_value = mean(total_pr))

New_mean <-
  mean_by_group |>
  filter(cond == "new")

Used_mean <-
  mean_by_group |>
  filter(cond == "used")


print(mean_by_group)
```

```
# A tibble: 2 × 2
  cond  mean_value
  <fct>      <dbl>
1 new         53.8
2 used        42.9
```

#This is our T- Test for Mario Kart

```
mario_kart_t_test<-
t.test(df = 139, total_pr ~ cond, data = mario_kart_selected, alternative

mario_kart_t_test
```

```
    Welch Two Sample t-test

data:  total_pr by cond
t = 8.64, df = 124, p-value = 2.3e-14
alternative hypothesis: true difference in means between group new and
group used is not equal to 0
95 percent confidence interval:
  8.4028 13.3963
sample estimates:
 mean in group new mean in group used
            53.771              42.871
```

```
result_table <- data.frame(
  new_mean = New_mean$mean_value,
  used_mean = Used_mean$mean_value,
  t_Statistic = mario_kart_t_test$statistic,
  Degrees_of_Freedom = mario_kart_t_test$parameter,
  P_Value = mario_kart_t_test$p.value
)

result_table_formatted <- kable(result_table, caption = "Mario Kart T-Test

result_table_formatted
```

Mario Kart T-Test Results by Used and New

	new_mean	used_mean	t_Statistic	Degrees_of_Freedom	P_Value
t	53.771	42.871	8.6406	123.99	0

#Now we make a table

```
filter(mario_kart_selected, cond %in% c("new", "used"))|>
  group_by(cond) |>
  summarise(mean = mean(total_pr),
            sd = sd(total_pr),
            median = median(total_pr),
            min = min(total_pr),
            max = max(total_pr),
            count = n()) |>
  kable(digits = 1)
```

cond	mean	sd	median	min	max	count
new	53.8	7.4	54.0	40.1	75	59
used	42.9	7.3	42.4	29.0	65	82