Mario_kart_analysis

We're importing the dataset from the Mario Kart

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

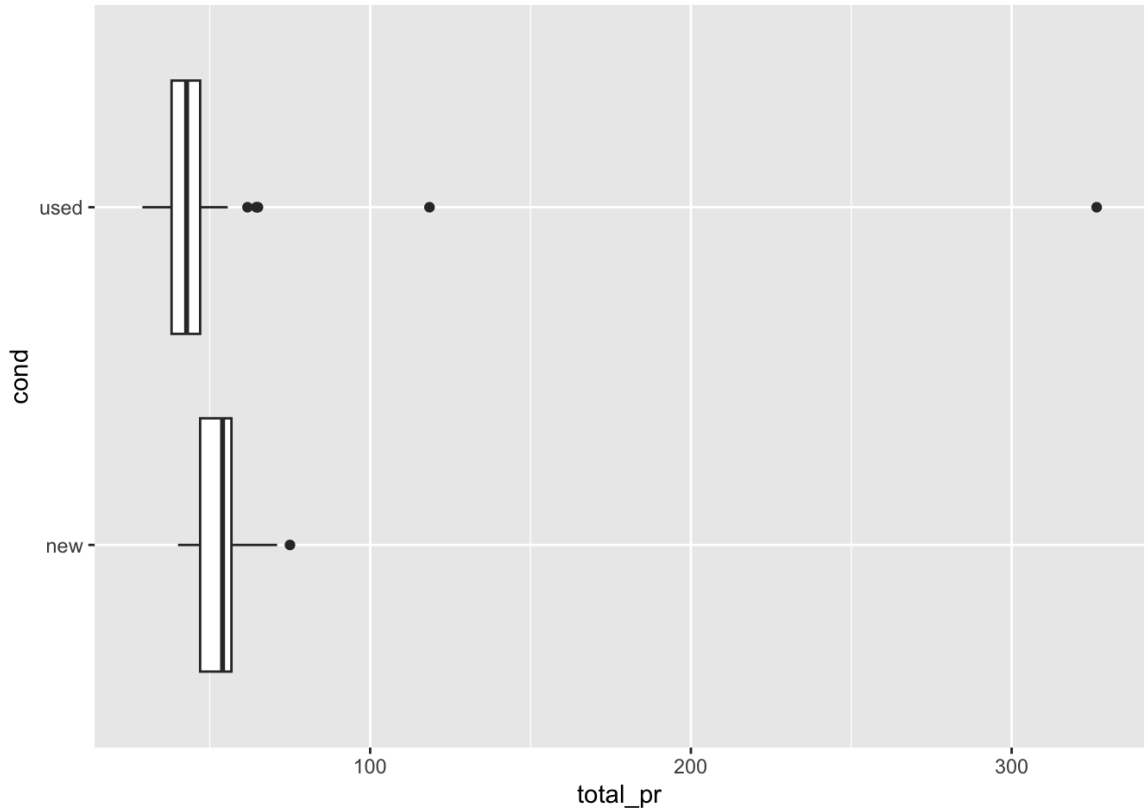
```
── Attaching core tidyverse packages ──────────────────────── tidyverse
2.0.0 ──
✔ dplyr     1.1.3     ✔ readr     2.1.4
✔ forcats   1.0.0     ✔ stringr   1.5.0
✔ ggplot2   3.4.3     ✔ tibble    3.2.1
✔ lubridate 1.9.2     ✔ tidyr     1.3.0
✔ purrr     1.0.2
── Conflicts ────────────────────────────────────────── 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()
```



#Now we have to remove the outliers

```
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, 118.5,65.02,64.50)))

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

```
new_mariokart <-
  mario_kart_selected |>
  filter(cond== "new")
  filter(new_mariokart, !(total_pr %in% c(75)))
```

```
# A tibble: 58 × 3
             id cond  total_pr
          <dbl> <fct>    <dbl>
 1 170392227765 new       71
 2 400076879560 new       66.4
 3 350261958546 new       65.0
 4 400077480990 new       65.0
 5 390103890073 new       65.0
 6 160366410200 new       64
 7 290356835914 new       64.0
 8 170390813853 new       63.5
 9 180416291487 new       62.9
10 170391971205 new       61
# ℹ 48 more rows
```

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

#Here we removed the outliers

```
mario_kart_selected <-
  filter(mario_kart_selected, !(total_pr %in% c(75)))

mario_kart_selected <-
    filter(mario_kart_selected, !(total_pr %in% c(326.51, 118.5, 65.02, 64
```

#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         53.4
2 used        42.3
```

#This is our T- Test for Mario Kart

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

mario_kart_t_test
```

```
    Welch Two Sample t-test

data:  total_pr by cond
t = 9.49, df = 118, p-value = 3.2e-16
alternative hypothesis: true difference in means between group new and
group used is not equal to 0
95 percent confidence interval:
  8.7687 13.3929
sample estimates:
 mean in group new mean in group used
           53.405             42.324
```

```
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	53.405	42.324	9.4901	118.35	0