1. A graph of a number of bars

   Description automatically generated with medium confidenceThe below histogram shows SPEED faceted by DISTANCE. Compare the distribution of SPEED for the 10k races with the distribution for the 20k races

Both distributions are left skewed, but the 20k is more skewed than the 10k. Additionally, the counts for the 20k are higher than for the 10k.

1. Perform a t.test to determine if there is a significant difference in the mean speed for each distance.

H0: mean 10k speed = mean 20k speed Ha: mean 10k speed mean 20k speed

**t =** -0.39936 **p-value =** 0.69; 0.69 > 0.05 **Conclusion:** Do not reject H0; We do not have significant evidence that the mean speed for 10k walkers is significantly different than the mean speed for 20k walkers.

```{r}

t.test(SPEED~DISTANCE, var.equal=TRUE, data=racewalking\_df)

```

* 1. A graph with blue and black squares

     Description automatically generatedBased on the below boxplot would you say that there is a significant difference in the mean speed for each distance and for each gender?

There looks to be a considerable difference in the mean speed for each gender and a slight difference in the mean speed for each distance.

1. After performing all of these tests on different models, are there any noticeable differences in the different races the athletes competed in? Are there any discernable patterns with records?

It seems that the main difference for walkers’ average speed in events is based on gender whereas distance does not seem to have as big of an impact on their speed. However, with records it does seem that both distance and gender factor in.